National College of Business Administration and Economics Lahore

SOME CONTRIBUTIONS TO SCRAMBLED RANDOMIZED RESPONSE TECHNIQUE USING AUXILIARY VARIABLE FOR PARAMETERS IN FINITE POPULATION

BY

IRAM SALEEM

DOCTOR OF PHILOSOPHY
IN
STATISTICS

NOVEMBER, 2017

NATIONAL COLLEGE OF BUSINESS ADMINISTRATION AND ECONOMICS

SOME CONTRIBUTIONS TO SCRAMBLED RANDOMIZED RESPONSE TECHNIQUE USING AUXILIARY VARIABLE FOR PARAMETERS IN FINITE POPULATION

BY

IRAM SALEEM

**A dissertation submitted to
Faculty of Social Sciences**

**In Partial Fulfillment of the
Requirements for the Degree of**

**DOCTOR OF PHILOSOPHY
IN
STATISTICS**

November, 2017

بسم الله الرحمن الرحيم

In the name of ALLAH,
The Most Beneficial,
The Most Merciful,

AUTHOR'S DECLARATION

I, **Iram Saleem** hereby state that my PhD thesis titled **"Some Contributions to Scrambled Randomized Response Technique using Auxiliary Variable for Parameters in Finite Population"** is my own work and has not been submitted previously by me for taking any degree from this university, **National College of Business Administration and Economics, Lahore** or anywhere else in the country/world.

At any time if my statement is found to be incorrect even after my graduate the university has the right to withdraw my PhD degree.

IRAM SALEEM
November, 2017

PLAGIARISM UNDERTAKING

I solemnly declare that research work presented in the thesis titled **“Some Contributions to Scrambled Randomized Response Technique using Auxiliary Variable for Parameters in Finite Population”** is solely my research work with no significant contribution from any other person. Small contribution/help whenever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero-tolerance policy of the HEC and **National College of Business Administration and Economics, Lahore** towards plagiarism. Therefore, I as an Author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly/cited.

I undertake that if I am found guilty of any formal plagiarism in the above title thesis even after award of PhD degree, the university reserves the right to withdraw/revoke my PhD degree and that HEC and the University has the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized thesis.

____

Iram Saleem

CERTIFICATE OF APPROVAL

This is to certify that research work presented in the thesis, entitled **"Some Contributions to Scrambled Randomized Response Technique using Auxiliary Variable for Parameters in Finite Population"** was conducted by **Iram Saleem** under the supervision of **Dr. Muhammad Hanif**.

No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the **Faculty of Social Sciences** in partial fulfillment of requirements for the degree of requirements for the degree of Doctor of Philosophy in the field of **Statistics**, School of Business Administration, National College of Business Administration and Economics, Lahore.

Student Name: **Iram Saleem** Signature: ______________

Examination Committee:

a) External Examiner 1:
Dr. Zahoor Ahmad Signature: ______________

b) External Examiner 2:
Dr. Aamir Sanaullah Signature: ______________
Associate Professor
Comsats Institute of Information Technology, Lahore

c) Internal Examiner 1:
Signature: ______________

Supervisor Name: **Dr. Muhammad Hanif** Signature: ______________

Co-Supervisor Name: **Dr. Muhammad Noor ul Amin** Signature: ______________

Name of Dean/HOD: **Dr. Abdul Khalique** Signature: ______________

Name of Rector: **Prof. Dr. Munir Ahmad** Signature: ______________

Dedicated
to

My beloved

Uncle Azeem Elahi

and

My Parents

ACKNOWLEDGEMENT

A research project like this depends largely on the contributions of many people, in their different ways. I would like to express my gratitude especially for the following:

Foremost, thank God for the wisdom and perseverance that He has bestowed upon me during this thesis task, and indeed, throughout my life:

‘*I can do everything through Him who gives me strength*.’
(Philippians 4:13).

To the National College of Business Administration and Economics (NCBA&E), I would like to express my sincere gratitude for providing me a conducive environment for completing this research work effectively.

I would like to offer my sincerest gratitude to my advisor and supervisor of this study Dr. Muhammad Hanif for the continuous support of my Ph.D. study and research. I salute him for his patience, enthusiasm and immense knowledge. It was a real privilege and honor for me to share his knowledge.

I am extremely grateful to my Co-supervisor Dr. Aamir Sanaullah for his constant support, availability and constructive suggestions, which helped me to complete this work.

I also gratefully acknowledge to the University of Waterloo, Canada for providing me environment and resources to improve my research work. Special thanks to Dr. Mary E. Thompson, for her encouragement, insightful comments and valuable help whenever I needed.

My sincere thanks also go to my fellow mates Uzma Yasmeen and Sadia Sajjad for the stimulating discussions.

Last but not the least; I would like to thank my parents and siblings for their unconditional support, encouragement, and love, without which I would not have come this far.

SUMMARY

In this dissertation, some generalized scrambled randomized response models for sensitive characteristics have been proposed using two scrambling variables. Similarly, some regression, ratio, exponential, regression-cum-exponential, regression-cum-ratio and ratio in exponential estimators have been proposed to estimate population characteristics for sensitive surveys using auxiliary information. The estimators are proposed for simple random sampling design for both single-phase sampling and two-phase sampling.

In Chapter 1, the discussion on sensitive surveys has been made. The techniques to collect information on sensitive characteristics such as randomized response and scrambled randomized response techniques have been introduced into more detail. Further, the use of auxiliary information and two-phase sampling have been illustrated. In Chapter 2, the review of literature regarding the use of auxiliary information in single-phase sampling and two-phase sampling have been discussed with the estimators developed by different statisticians for both sensitive and non-sensitive surveys. Various development on randomized response models in literature have also been presented.

The major work of this dissertation start from Chapter 3. In this chapter, four generalized scrambled randomized response models have been proposed combining additive and multiplicative models. These four generalized models have been proposed using two scrambling variables with known distribution. The expressions of the mean, variance, covariance and correlation have been derived for each of the proposed models. Additionally, the privacy measure have been derived for some existing models presented in literature review and the proposed models. The privacy protection comparisons between existing models and proposed models have also been discussed.

In Chapter 4, the generalized exponential-type estimators have been constructed using two auxiliary variables to estimate population mean of the sensitive variable. The bias and mean square error have been derived for each proposed estimator. To examine the performance of the proposed generalized estimators, the simulation study have been performed under the observed response using additive and proposed scrambled randomized response models.

In Chapter 5, the regression, ratio, regression-cum-ratio, regression-cum exponential and ratio in exponential-type estimators have been proposed under two-phase sampling to estimator population mean of sensitive study variable. The estimators have been proposed for three cases of two-phase sampling such as full-information-case, partial-information-case and no-information-case.

The expressions of the bias and mean square error have been derived for each proposed estimator. Additionally, the simulation study has been conducted to examine the performance of estimator using additive and proposed models.

To estimate the population variance of sensitive study variable, some exponential estimators have been proposed in Chapter 6. These estimators have been presented for both single-phase and two-phase sampling. In this chapter, the additive model is considered to estimate population variance. The expressions of the bias and mean square error have been derived. The simulation study have also been presented for both single-phase and two phase sampling.

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 SENSITIVE SURVEYS

In a survey, the aim is to gather information on the issues of the population under study. So, the survey statisticians are interested to bring out information through different modes of data collection from some or all the units selected in the sample.

As we know, it is not possible to select a complete sample, to design an ideal questionnaire, or to get complete response from survey interviews. Because of this, errors can be introduced in surveys. The survey errors may be due to sampling or non-sampling errors. *Sampling errors* occur if there is variation in the representativeness of the sample and these errors may be minimized by increasing the size of the sample size. *Non-sampling errors* arise due to many reasons such as coverage, processing, measurement or non-response etc. The *coverage error* occurs when some elements of the population are excluded or included or misclassified mistakenly in the sampling frame. This error causes the bias for the estimation process. By adapting different data collection mode the coverage error may be reduced. The *processing error* follows during imputation or coding mechanism of data and may cause bias or increase the variance of the estimates. The processing errors may be controlled by using quality control techniques. The *measurement error* may be produced by the respondent selected in the sample or by the interviewer or the questionnaire is not understandable by the respondent. So the measurement error may be defined as the difference between the recorded response to a particular question asked and the actual value. The measurement error may only be minimized if the questionnaire design is made understandable for the respondents or trained interviewer is appointed for data collection.

The most important error that cannot be neglected is the *non-response error*. This may occur if the respondent is not available or unable to answer or refuse to provide answer to the required information. In different fields of research, direct questioning is often used to collect required information which may become a major cause of the refusals from the respondents. Sometimes the data collection might be of characteristics which are illegal or personal in nature such as criminal behavior, drunken driving, reuse of infected syringes, carrying an HIV virus, etc. and the questions on such characteristics are known to be threatening or sensitive questions. Asking questions on sensitive issues is about requesting the respondent to provide information they are not willing to answer. In such complex surveys, two types of behavior may be expected from respondents to answer; non-response or response error. Mostly, non-response may occur due to the high rate of refusals to respond to invasive concerns (Tourangeau and Smith 1996). There are two natures of non-response may be faced in the survey. If the respondent refuses to answer some specific questions, such a non-response is said to be item non-response. On the other hand, there is a situation where the respondent refuses to participate in the survey, known as unit non-response. On the contrary, response error or misreporting arises where the respondent agrees to answer but he/she may provide incorrect or false feedback. These errors commonly cause a large systematic bias to estimate population parameter on sensitive questions.

We cannot avoid non-response errors but can provide anonymity, confidentiality and privacy to the respondents which may lead to minimize such errors (Bradburn, Sudman and Wansink, 2004). A good survey method should guarantee the privacy and confidentiality of the respondents to obtain accurate reporting and reduce the non-response or response error. The term 'confidentiality' refers to the stability of the data collected through the survey researcher. The intention of the research study is how the confidentiality is respected about the private concerns of the participants. On the other hand, 'anonymity' introduced the elimination of participant's identity. The term 'privacy, is concerned with the response provided by the participant within the community. So, by securing the anonymity and privacy, confidentiality may be protected.

1.2 RANDOMIZED RESPONSE MODEL

Randomized response model by Warner (1965). The method includes using a device to maximize the anonymity and privacy of response and reduce systematic bias in complex surveys enclosing illegal behavior known as randomized response technique (*RRT*). In *RRT*, the respondent has a choice to answer either a sensitive or a non-sensitive question. "*For example, 'In the past five years have you used a needle to inject illegal drugs?' or 'Is your birthday in August?' The respondent is then told to use randomization device, such as flipping a coin, to decide which question to answer- if 'heads', to answer the first question and if 'tails', the second. The interviewer or test administrator will neither see nor be told the results of the coin toss and, hence, which question the respondent was supposed to answer. Researchers, however, would be able to compute the probability of a particular (yes) response to the non-threating question. Any departure from this probability in* *the actual responses can be used to estimate the proportion of responses to the threatening item*".

Here, the interviewer is unaware of the question selected as well as the answer given by the respondent. Through this process, hopefully the respondent feels comfortable to provide a trustworthy response. The procedure can be applied on dichotomous surveys. Greenberg et al. (1967) extended the work of Warner (1965) to tri-chotomous case to estimate the proportion of three mutually exclusive groups. Ericksson (1973) also followed Warner's method and used a single randomization device to estimate multinomial proportions for sensitive problems. Pollock and Bek (1976) made comparison among three models i.e. Greenberg (1969), additive and multiplicative models and many other also contributed on *RRT* for different situations.

1.3 SCRAMBLED RANDOMIZED RESPONSE TECHNIQUE

The above method is useful only for qualitative sensitive surveys. Greenberg et al. (1971) made an extension on the work of Warner (1965) and presented an alternative method for quantitative data named as "*unrelated question randomized response model*". According to Greenberg et al. (1971) method, the respondent is asked to choose the sensitive question with preassigned probability *p* and another inoffensive question with probability 1-*p*. This method is not much useful in practice.

An alternative method was introduced by Eichhorn and Hayre (1983) known as the Scrambled Randomized Response (*SRR*) method. This method involves the respondent quantitative reply of a sensitive question asked, which is then multiplied by a random number from a known distribution generated by the respondents themselves following some supplied mechanism. The interviewer fairly obtains the resulted product and does not know the random number used for scrambling the true response.

The *SRR* method is a special case of Warner (1971) and Pollock and Bek (1976). The *SRR* models can be categorized as: a *forced SRR* model, a *partial SRR* model or an *optional SRR* model. In the *forced SRR* method, the respondent is forced by the randomization device to report to the scrambled response (Hussain and Shabbir 2007). The *partial SRR* method involves a known proportion of participants to report the scrambled response by using one or more randomization devices (Bar-Lev et al., 2004; Ryu et al., 2005). In an *optional SRR* method, the researcher provides an option to the participants to report either true (non-sensitive variable) response or scrambled (sensitive variable) response to the question asked.

Further, the *SRR* models can be classified as multiplicative scrambling model (Eichhorn and Hayre, 1983), additive scrambling model (Himmelfarb and Edgel, 1980), subtractive scrambling model (Hussain, 2012), or mixture of additive and multiplicative scrambling models (Huang, 2010).

1.4 AUXILIARY INFORMATION AND TWO-PHASE SAMPLING

To obtain high precision for estimating population characteristics, one of the two methods may be applied. The first method is to use adequate sampling design and the second one is to use auxiliary information. The auxiliary information is supplementary information, correlated with the study variable and is suitable to improve the efficiency of the variable under the investigation. The use of auxiliary information can provide precise estimates at both estimation and design stage. Grant (1662) first used auxiliary information to estimate the population of England. Likewise, Laplace (1820) applied auxiliary information to estimation the population of France.

In many surveys, auxiliary information is readily available. For such cases, single-phase sampling is said to be more effective to use. But sometimes, the auxiliary information is not readily available or when the cost of drawing large sample is too high. Neyman (1938) introduced the concept of two-phase sampling. In two-phase sampling, a comparatively large sample is selected from the population at first-phase and auxiliary variable is obtained from this first sample. Then a second sample is selected from the first sample, known as second phase sample. The study variable y is obtained from second phase sample.

A ratio estimator is used if positive and linear correlation exits between auxiliary variable and the study variable [Cochran 1940], and product estimator is used if the linear correlation is negative [Robson 1957]. Occasionally, this correlation is not strong enough to use usual ratio or product estimator, here in such situation exponential estimators may be used alternatively.

Srivastava (1971) and Bhal and Tuteja (1991) developed some exponential type estimators for single-phase sampling. Following Bhal and Tuteja (1991) many survey statisticians extended the work of exponential estimators in single-phase sampling to two-phase sampling, such as Singh and Vishwakarma (2007), and Sanaullah et al. (2012).

1.5 STATEMENT OF THE PROBLEM

From the available literature, it is observed that a limited work is available on the estimation of population mean for sensitive characteristic. Particularly, a few research works is available in literature on the estimation of population variance using *RRT*. We have noticed no contribution to the *RR* models on utilizing more than one auxiliary variable however *RRT* by different researcher have been proposed using single auxiliary variables. Also, we have noticed that in literature only linear and positive relation between the sensitive study variable and auxiliary variable are of more interest but in many situation this linear relation may not be true. In such a situation we are to see some more non-linear relationships. In this study, our task is to study some more situations of non-linear relationships between the sensitive study variable and the non-sensitive auxiliary variable(s).

1.6 OBJECTIVE OF STUDY

i) One of our major objectives is to develop some improved *SRR* model for the estimation of population mean and variance of the sensitive variable of interest.

ii) Another major objective is to develop some new estimators to estimate the population characteristics for scramble randomize response technique using more than one non-sensitive auxiliary variable. These estimators will be proposed if the relationship between the study and auxiliary variables is linear and the relationship is of exponential nature.

iii) Conduct simulation study using artificial data to examine the performance of estimators using *SRR* models.

1.7 NOTATIONS

1.7.1 Notations for the Estimation of Population Mean of Non-Sensitive Study Variable

Consider a finite population consists of N units $U=\{U_1, U_2, \ldots, U_N\}$. Let y be the non-sensitive variable under study with population mean and variance respectively as $\bar{Y}=\frac{1}{N}\sum_{i=1}^{N}Y_i$ and $S_y^2=\frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i-\bar{Y}\right)^2$. Let x_k be the

kth $\left(k=1,2,..,q\right)$ non-sensitive auxiliary variable having population mean and variance as $\bar{X}_k=\frac{1}{N}\sum_{i=1}^{N}X_k$ and $S_{x_k}^2=\frac{1}{N-1}\sum_{i=1}^{N}\left(X_{ki}-\bar{X}_k\right)^2$. Let a sample of size n_1 is selected from the population through *SRSWOR* and consider $\bar{y}=\frac{1}{n_1}\sum_{i=1}^{n_1}y_i$ and $\bar{x}_k=\frac{1}{n_1}\sum_{i=1}^{n_1}x_k$ be the sample means for y and x_k respectively. The population coefficient of variation for the study variable and the auxiliary variable are respectively as $C_y^2=\frac{S_y^2}{\bar{Y}^2}$ and $C_{x_k}^2=\frac{S_{x_k}^2}{\bar{X}_k^2}$. Also, the terms of correlation and covariance between y and x_k are correspondingly as $\rho_{yx_k}=\frac{S_{yx_k}}{S_{x_k}S_y}$ and $S_{yx_k}^2=\frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i-\bar{Y}\right)\left(X_{ki}-\bar{X}_k\right)$.

Further we consider $e_{\bar{y}}=\frac{\bar{y}-\bar{Y}}{\bar{Y}}$ and $e_{\bar{x}k}=\frac{\bar{x}_k-\bar{X}_k}{\bar{X}_k}$ are the sampling errors respectively on the study variable and the k^{th} auxiliary variables such that

$$E_D\left(e_{\bar{y}}\right)=E_D\left(e_{\bar{x}k}\right)=0, \tag{1.7.1}$$

$$\left.\begin{aligned}
E_D\left(e_{\bar{y}}\right)^2&=E_D\left(\frac{\bar{y}-\bar{Y}}{\bar{Y}}\right)^2=\frac{1}{\bar{Y}^2}Var\left(\bar{y}\right)=\theta_1C_y^2\\
E_D\left(e_{\bar{x}_k}\right)^2&=E_D\left(\frac{\bar{x}_k-\bar{X}_k}{\bar{X}_k}\right)^2=\frac{1}{\bar{X}_k{}^2}Var\left(\bar{x}_k\right)=\theta_1C_{x_k}^2\\
E_D\left(e_{\bar{y}}e_{\bar{x}_k}\right)&=E_D\left[\left(\frac{\bar{y}-\bar{Y}}{\bar{Y}}\right)\left(\frac{\bar{x}_k-\bar{X}_k}{\bar{X}_k}\right)\right]=\frac{1}{\bar{Y}\bar{X}_k}Cov\left(\bar{y},\bar{x}_k\right)=\theta_1\rho_{yx_k}C_yC_{x_k}
\end{aligned}\right\}, \tag{1.7.2}$$

where $\theta_1=\left(\frac{1}{n_1}-\frac{1}{N}\right)$.

Let another sample of size n_2 from the first sample n_1 is selected using *SRSWOR* and the information for the study variable y along with auxiliary variable(s) say x_1 or x_2 or both (if necessary) are obtained. This is the second-phase sample and the notations for three different cases in two-phase sampling are defined as,

I. When complete information on the required parameters of auxiliary variable x_1 and x_2 is readily available. This case is said to be full-information case (*FIC*). Assume that $\bar{x}'_k = \frac{1}{n_1}\sum_{i=1}^{n_1} x'_{ki}$ are the sample mean of auxiliary variables x_k obtained from the first-phase sample of size n_1 and $\bar{y}'' = \frac{1}{n_2}\sum_{i=1}^{n_2} y''_i$ is the sample mean of the study variable *y* from the second-phase sample of size n_2. Furthermore, let us suppose $e''_{\bar{y}} = \frac{\bar{y}'' - \bar{Y}}{\bar{Y}}$, $e'_{\bar{x}1} = \frac{\bar{x}'_1 - \bar{X}_1}{\bar{X}_1}$ and $e'_{\bar{x}2} = \frac{\bar{x}'_2 - \bar{X}_2}{\bar{X}_2}$ are the sampling errors on the study variable and the auxiliary variables for full-information case such that,

$$E_D\left(e''_{\bar{y}}\right) = E_D\left(e'_{\bar{x}k}\right) = E_D\left(e'_{\bar{x}j}\right) = 0, \tag{1.7.3}$$

$$E_D\left(e''_{\bar{y}}\right)^2 = E_D\left(\frac{\bar{y}'' - \bar{Y}}{\bar{Y}}\right)^2 = \frac{1}{\bar{Y}^2}Var\left(\bar{y}''\right) = \theta_2 C_y^2, \tag{1.7.4}$$

$$\left.\begin{aligned}
E_D\left(e'_{\bar{x}_k}\right)^2 &= E_D\left(\frac{\bar{x}'_k - \bar{X}_k}{\bar{X}_k}\right)^2 = \frac{1}{\bar{X}_k{}^2}Var\left(\bar{x}_k\right) = \theta_1 C_{x_k}^2 \\
E_D\left(e''_{\bar{y}} e'_{\bar{x}_k}\right) &= E_D\left[\left(\frac{\bar{y}'' - \bar{Y}}{\bar{Y}}\right)\left(\frac{\bar{x}'_k - \bar{X}_k}{\bar{X}_k}\right)\right] \\
&= \frac{1}{\bar{Y}\bar{X}_k} Cov\left(\bar{y}'', \bar{x}'_k\right) = \theta_1 \rho_{yx_k} C_y C_{x_k} \\
E_D\left(e'_{\bar{x}_k} e'_{\bar{x}_j}\right) &= E_D\left[\left(\frac{\bar{x}'_k - \bar{X}_k}{\bar{X}_k}\right)\left(\frac{\bar{x}'_j - \bar{X}_j}{\bar{X}_j}\right)\right] \\
&= \frac{1}{\bar{X}_k \bar{X}_j} Cov\left(\bar{x}'_k, \bar{x}'_j\right) = \theta_1 \rho_{x_k x_j} C_{x_k} C_{x_j} \quad (k \neq j)
\end{aligned}\right\}, \tag{1.7.5}$$

where $\theta_2 = \left(\frac{1}{n_2} - \frac{1}{n_1}\right)$ and $\theta_3 = \theta_2 - \theta_1$.

II. When information on the required parameters of any auxiliary variable say x_k is completely available and the information on x_j is not. This case is considered as partial-information case (*PIC*). Consider $\bar{x}''_j = \frac{1}{n_2}\sum_{i=1}^{n_2} x''_{ji}$ be the second-phase sample mean of the auxiliary variable x_j, assume $e''_{\bar{x}j} = \frac{\bar{x}''_j - \bar{X}_j}{\bar{X}_j}$ be the sampling error revised under partial-information case on the auxiliary variable x_j such that,

$$E_D\left(e''_{\bar{x}j}\right) = 0, \tag{1.7.6}$$

$$\left.\begin{aligned}
E_D\left(e''_{\bar{x}_j}\right)^2 &= E_D\left(\frac{\bar{x}''_j - \bar{X}_j}{\bar{X}_j}\right)^2 = \frac{1}{\bar{X}_j^{\,2}} Var\left(\bar{x}_j\right) = \theta_2 C^2_{x_j} \\
E_D\left(e''_{\bar{y}} e''_{\bar{x}_j}\right) &= E_D\left[\left(\frac{\bar{y}'' - \bar{Y}}{\bar{Y}}\right)\left(\frac{\bar{x}''_j - \bar{X}_j}{\bar{X}_j}\right)\right] \\
&= \frac{1}{\bar{Y}\bar{X}_j} Cov\left(\bar{y}'', \bar{x}''_j\right) = \theta_2 \rho_{yx_j} C_y C_{x_j} \\
E_D\left(e'_{\bar{x}_k} e''_{\bar{x}_j}\right) &= E_D\left[\left(\frac{\bar{x}'_k - \bar{X}_k}{\bar{X}_k}\right)\left(\frac{\bar{x}''_j - \bar{X}_j}{\bar{X}_j}\right)\right] \\
&= \frac{1}{\bar{X}_k \bar{X}_j} Cov\left(\bar{x}'_k, \bar{x}''_j\right) = \theta_1 \rho_{x_k x_j} C_{x_k} C_{x_j}
\end{aligned}\right\}, \tag{1.7.7}$$

III. When there is no complete information available on both auxiliary variables (x_k, x_j), such a case is considered to be no-information case (*NIC*). Under this case the sample mean of auxiliary variable x_k is as $\bar{x}''_k = \frac{1}{n_2}\sum_{i=1}^{n_2} x''_{ki}$ and also assuming $e''_{\bar{x}k} = \frac{\bar{x}''_k - \bar{X}_k}{\bar{X}_k}$ as the sampling error term on *k*th auxiliary variable as

$$E_D\left(e''_{\bar{x}k}\right) = 0, \tag{1.7.8}$$

$$\left.\begin{aligned}
E_D\left(e''_{\bar{x}_k}\right)^2 &= E_D\left(\frac{\bar{x}''_k-\bar{X}_k}{\bar{X}_k}\right)^2 = \frac{1}{\bar{X}_k{}^2}Var\left(\bar{x}_k\right)=\theta_2 C_{x_k}^2\\
E_D\left(e''_{\bar{y}}e''_{\bar{x}_k}\right) &= E_D\left[\left(\frac{\bar{y}''-\bar{Y}}{\bar{Y}}\right)\left(\frac{\bar{x}''_k-\bar{X}_k}{\bar{X}_k}\right)\right]\\
&= \frac{1}{\bar{Y}\bar{X}_k}Cov\left(\bar{y}'',\bar{x}''_k\right)=\theta_2\rho_{yx_k}C_y C_{x_k}\\
E_D\left(e''_{\bar{x}_k}e''_{\bar{x}_j}\right) &= E_D\left[\left(\frac{\bar{x}''_k-\bar{X}_k}{\bar{X}_k}\right)\left(\frac{\bar{x}''_j-\bar{X}_j}{\bar{X}_j}\right)\right]\\
&= \frac{1}{\bar{X}_k\bar{X}_j}Cov\left(\bar{x}''_k,\bar{x}''_j\right)=\theta_2\rho_{x_kx_j}C_{x_k}C_{x_j}
\end{aligned}\right\}. \qquad (1.7.9)$$

1.7.2 Notations for the Estimation of Population Variance of Non-Sensitive Study Variable

Let us consider the sample variances of *y* and x_1 from the first sample as $s_y^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}\left(y_i-\bar{y}\right)^2$ and $s_{x1}^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}\left(x_{1i}-\bar{x}_1\right)^2$ respectively. Also, assume that $e_{s_y^2}=\frac{s_y^2-S_y^2}{S_y^2}$ and $e_{s_{x_1}^2}=\frac{s_{x_1}^2-S_{x_1}^2}{S_{x_1}^2}$ are the sampling error terms on the study variable and the auxiliary variable such that o the valuable notations are defined below,

$$E_D\left(e_{s_y^2}\right)=E_D\left(e_{s_{x_1}^2}\right)=0\,, \qquad (1.7.10)$$

$$\left.\begin{aligned}
E_D\left(e_{s_y^2}\right)^2 &= E_D\left(\frac{s_y^2 - S_y^2}{S_y^2}\right)^2 = \theta_1\delta_{40} \\
E_D\left(e_{s_{x_1}^2}\right)^2 &= E_D\left(\frac{s_{x_1}^2 - S_{x_1}^2}{S_{x_1}^2}\right)^2 = \theta_1\delta_{04} \\
E_D\left(e_{s_y^2}e_{s_{x_1}^2}\right) &= E_D\left[\left(\frac{s_y^2 - S_y^2}{S_y^2}\right)\left(\frac{s_{x_1}^2 - S_{x_1}^2}{S_{x_1}^2}\right)\right] = \theta_1\delta_{22}
\end{aligned}\right\}, \qquad (1.7.11)$$

where $\delta_{ab} = \frac{1}{(N-1)}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^a\left(X_i - \bar{X}\right)^b$ for (a,b) =0,1,2,3,4.

1.7.3 Notations to Estimate Population Characteristics for Sensitive Variable of Interest

i. Let z be the coded response to estimate sensitive variable y with sample mean $\bar{z} = \frac{1}{n_1}\sum_{i=1}^{n_1} z_i$ corresponding to the population mean $\bar{Z} = \frac{1}{N}\sum_{i=1}^{N} Z_i$ and sample variance $s_z^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}\left(z_i - \bar{z}\right)^2$ for corresponding to the population variance as $S_Z^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(Z_i - \bar{Z}\right)^2$ respectively. Here x_k are observed directly but y will be obtained in scrambled version. The population coefficient of variation for the coded response z is as $C_Z^2 = \frac{S_Z^2}{\bar{Z}^2}$ and the correlation and covariance between z and x_k are correspondingly as $\rho_{zx_k} = \frac{S_{zx_k}}{S_{x_k}S_z}$ and $S_{zx_k}^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(Z_i - \bar{Z}\right)\left(X_{ki} - \bar{X}_k\right)$. Let S and R be the two positive independent scrambling variables and are normally distributed with means as $\bar{S} = \frac{1}{N}\sum_{i=1}^{N} S_i, \bar{R} = \frac{1}{N}\sum_{i=1}^{N} R_i$ and variances $S_S^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(S_i - \bar{S}\right)^2$ and $S_R^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(R_i - \bar{R}\right)^2$ respectively.

Moreover, the sampling error for the coded response z to estimate the population mean is $e_{\bar{z}} = \frac{\bar{z} - \bar{Z}}{\bar{Z}}$ such that

$$E_D\left(e_{\bar{z}}\right) = 0. \tag{1.7.12}$$

$$\left.\begin{aligned} E_D\left(e_{\bar{z}}\right)^2 &= E_D\left(\frac{\bar{z}-\bar{Z}}{\bar{Z}}\right)^2 = \frac{1}{\bar{Z}^2}Var\left(\bar{z}\right) = \theta_1 C_z^2 \\ E_D\left(e_{\bar{z}} e_{\bar{x}_k}\right) &= E_D\left[\left(\frac{\bar{z}-\bar{Z}}{\bar{Z}}\right)\left(\frac{\bar{x}_k-\bar{X}_k}{\bar{X}_k}\right)\right] \\ &= \frac{1}{\bar{Z}\bar{X}_k}Cov\left(\bar{z},\bar{x}_k\right) = \theta_1 \rho_{zx_k} C_z C_{x_k} = \theta_1 C_{zx_k} \end{aligned}\right\}. \tag{1.7.13}$$

ii. Similarly, to estimate population variance for sensitive study variable the sampling error for the coded response z is as $e_{s_z^2} = \frac{s_z^2 - S_z^2}{S_z^2}$ and the expectations are

$$\left.\begin{aligned} E_D\left(e_{s_z^2}\right)^2 &= E_D\left(\frac{s_Z^2 - S_Z^2}{S_Z^2}\right)^2 = \theta_1 \vartheta_{40} \\ E_D\left(e_{s_{x_1}^2}\right)^2 &= E_D\left(\frac{s_{x_1}^2 - S_{x_1}^2}{S_{x_1}^2}\right)^2 = \theta_1 \vartheta_{04} \\ E_D\left(e_{s_z^2} e_{s_{x_1}^2}\right) &= E_D\left[\left(\frac{s_Z^2 - S_Z^2}{S_Z^2}\right)\left(\frac{s_{x_1}^2 - S_{x_1}^2}{S_{x_1}^2}\right)\right] = \theta_1 \vartheta_{22} \end{aligned}\right\}, \tag{1.7.14}$$

where $\vartheta_{ab} = \frac{1}{(N-1)}\sum_{i=1}^{N}\left(Z_i - \bar{Z}\right)^a\left(X_i - \bar{X}\right)^b$, for (a,b) =0,1,2,3,4.

iii. Similarly, for two-phase sampling to estimate population mean of sensitive study variable y the sample mean for the coded response at second-phase is $\bar{z}'' = \frac{1}{n_2}\sum_{i=1}^{n_2} z_i''$. The error term on z for second phase is consider as $e''_{\bar{z}} = \frac{\bar{z}'' - \bar{Z}}{\bar{Z}}$, sample mean for the notations for two-phase replacing y to z are given as follow,

$$\left.\begin{aligned}
&E_D\left(e''_{\bar z}\right)^2 = E_D\left(\frac{\bar z'' - \bar Z}{\bar Z}\right)^2 = \frac{1}{\bar Z^2}Var\left(\bar z''\right) = \theta_2 C_z^2\\
&E_D\left(e''_{\bar z}e'_{\bar x_k}\right) = E_D\left[\left(\frac{\bar z'' - \bar Z}{\bar Z}\right)\left(\frac{\bar x'_k - \bar X_k}{\bar X_k}\right)\right]\\
&\qquad = \frac{1}{\bar Z \bar X_k}Cov\left(\bar z'', \bar x'_k\right) = \theta_1 \rho_{zx_k} C_z C_{x_k} = \theta_1 C_{zx_k}\\
&E_D\left(e''_{\bar z}e''_{\bar x_k}\right) = E_D\left[\left(\frac{\bar z'' - \bar Z}{\bar Z}\right)\left(\frac{\bar x''_k - \bar X_k}{\bar X_k}\right)\right]\\
&\qquad = \frac{1}{\bar Z \bar X_k}Cov\left(\bar z'', \bar x''_k\right) = \theta_2 \rho_{zx_k} C_z C_{x_k} = \theta_2 C_{zx_k}
\end{aligned}\right\}. \qquad (1.7.15)$$

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

In this chapter, a brief discussion on some existing estimators to estimate population mean for non-sensitive as well as sensitive study variable has been delivered along with the bias and the mean squared error. The chapter is divided as, section 2.1 discuss the different estimators available to estimate population mean of non-sensitive study variable using auxiliary information. In section 2.2, some available estimators on two-phase sampling has been presented with the bias and the mean square error. To deal with the situation of sensitive study variables some randomized response techniques are listed in section 2.3. In section 2.4, the work on the estimation of sensitive variable of interest using different response models are given with the bias and the mean square error. And section 2.5 presents some of the estimators available in literature for estimating population variable of sensitive and non-sensitive study variable.

2.2 ESTIMATORS FOR NON-SENSITIVE STUDY VARIABLE

Graunt (1662) used auxiliary information to estimate the population of England. Then, the population of France was estimated by Laplace (1951) utilizing auxiliary information. Neyman (1934) specified the foundation of sampling theory. Cochran (1940) introduced ratio estimator using auxiliary information by observing the correlation between study variable y and auxiliary variable x. The ratio estimator is given as:

$$t_1 = \bar{y}\left[\frac{\bar{X}}{\bar{x}}\right]. \tag{2.2.1}$$

The bias and the mean square error of the ratio estimator are as,

$$Bias(t_1) = \theta\bar{Y}C_x^2\left(1-\rho_{yx}\frac{C_y}{C_x}\right), \tag{2.2.2}$$

and

$$MSE(t_1) = \bar{Y}^2\theta\left[C_y^2 + C_x^2 - 2\rho_{yx}C_xC_y\right], \tag{2.2.3}$$

where $\theta=\left(\frac{1}{n}-\frac{1}{N}\right), C_y^2=\frac{S_y^2}{\bar{Y}^2}, C_x^2=\frac{S_x^2}{\bar{X}^2}$.

Robson (1957) developed product estimator under the condition relation is linear but when negative correlation is observed between the study and the auxiliary variable. The product estimator is as,

$$t_2=\bar{y}\left[\frac{\bar{x}}{\bar{X}}\right], \tag{2.2.4}$$

and the bias and the *MSE* of t_2 are given as,

$$Bias\left(t_2\right)=\theta\bar{Y}C_x^2\left(1+\rho_{yx}\frac{C_y}{C_x}\right), \tag{2.2.5}$$

and

$$MSE\left(t_2\right)=\bar{Y}^2\theta\left[C_y^2+C_x^2+2\rho_{yx}C_xC_y\right]. \tag{2.2.6}$$

So, the product estimator introduced by Robson (1957) considered to be more precise than usual mean and ratio estimator when the correlation among *y* and *x* is negative. Murthy (1964) revisited the ratio and product estimator as an unbiased estimator and also discussed their properties. He recommended that the ratio and product estimators performs more efficiently if $\rho_{yx}C_y/C_x>1/2$ and $\rho_{yx}C_y/C_x<-1/2$ holds for them respectively. Motivated from Cochran (1940) and Robson (1957), a new idea to generalize the estimators was presented by Srivastava (1967). He proposed a generalized estimator given as:

$$t_3=\bar{y}\left[\frac{\bar{X}}{\bar{x}}\right]^{\alpha}, \tag{2.2.7}$$

where $\alpha\epsilon\left[1,-1\right]$ is a suitable constant. If $\alpha=1$ the above estimator reduces to the ratio estimator presented by Cochran (1940) and if $\alpha=-1$ the estimator becomes Robson (1957) product estimator. The derived bias and the *MSE* of t_3 are as follows,

$$Bias\left(t_3\right)=\theta\bar{Y}\alpha C_x^2\left(\frac{\left(\alpha-1\right)}{2}-\rho_{yx}\frac{C_y}{C_x}\right), \tag{2.2.8}$$

and

$$MSE\left(t_3\right)=\bar{Y}^2\theta\left[C_y^2+\alpha^2C_x^2-2\alpha\rho_{yx}C_xC_y\right]. \tag{2.2.9}$$

Srivastava (1967) conferred that ratio and product estimators are useful under the conditions presented by Murthy (1964). The suggested estimator above is a good estimate if $\alpha = -\rho_{yx}\dfrac{C_y}{C_x}$ and has the minimum variance and the bias as compared to ratio and product estimators.

Further to estimate the population mean with unknown values of *y* Walsh (1970) introduced a generalized ratio estimator using known values of auxiliary variable *x*. The generalized family of estimator given as follows:

$$t_4 = \bar{y}\left[\frac{\bar{X}}{\alpha\bar{x} + (1-\alpha)\bar{X}}\right], \tag{2.2.10}$$

where α is a suitable constant. If $\alpha = 1$ the above estimator reduces to the ratio estimator presented by Cochran (1940). If $\alpha = 0$ the estimator becomes unbiased mean estimator. The bias and the *MSE* of t_4 are,

$$Bias(t_4) = \theta\bar{Y}\alpha C_x^2\left(\alpha - \rho_{yx}\frac{C_y}{C_x}\right), \tag{2.2.11}$$

and

$$MSE(t_4) = \bar{Y}^2\theta\left[C_y^2 + \alpha^2 C_x^2 - 2\alpha\rho_{yx}C_x C_y\right]. \tag{2.2.12}$$

The *MSE* is minimum for $\alpha = \rho_{yx}\dfrac{C_y}{C_x}$ and the variance of t_4 is as small as achievable for linear regression estimator.

Moreover, Khoshnevisan (2007) presented the more generalized form of Walsh's (1970) ratio-type estimator under *SRSWOR* scheme to estimate the population mean of study variable. The estimator is as,

$$t_5 = \bar{y}\left[\frac{a\bar{X} + b}{\alpha(a\bar{x} + b) + (1-\alpha)(a\bar{X} + b)}\right]^g, \tag{2.2.13}$$

where *g*, α, *a* and *b* are suitable constants. If $g = 1$ and $\alpha = 1$, the estimator becomes an unbiased mean estimator. If the choices are $g = -1, \alpha = 1, a = 1$ and b=0, the estimator becomes product estimator. If $g = 1, \alpha = 1, a = 1$ and b=0, the estimator t_5 reduces to ratio estimator. If $g = g, \alpha = 1, a = 1$ and *b=0,* the estimator t_5 becomes Srivastava (1967) generalized estimator. Also, if $g = 1, \alpha = \alpha, a = 1$ and *b*=0, the estimator t_5 reduces to Walsh (970) estimator. And if $g = 0$ the estimator t_5 becomes usual unbiased estimator. The bias and the *MSE* of t_5 are,

$$Bias(t_5) = \bar{Y}\theta\left[\frac{g(g+1)}{2}\alpha^2\tau^2C_x^2 - \alpha\tau g\rho_{yx}C_xC_y\right], \tag{2.2.14}$$

and

$$MSE(t_5) = \bar{Y}^2\theta\left[C_y^2 + g^2\alpha^2\tau^2C_x^2 - 2\alpha\tau g\rho_{yx}C_xC_y\right] \ , \tag{2.2.15}$$

where $\tau = \dfrac{a\bar{X}}{a\bar{X} + b}$ and the *MSE* is minimum for $\alpha = \rho_{yx}\dfrac{C_y}{C_x}\dfrac{1}{\tau\alpha}$. The resultant mean squared error obtained is the same as that of the approximate variance of the linear regression estimator. Furthermore, Gupta and Shabbir (2008), Koyuncu and Kadilar (2009), Solanki et al. (2012) and many other contributed in the development of different ratio and product type estimator to attain maximum efficiency.

Other than above mentioned ratio and product type estimators, Watson (1937) considered the situation if there exist a linear relationship between the variable *y* and variable *x* and the line does not pass through the origin then regression estimators are considered to be efficient. This regression estimator was revisited by Cochran (1940) to present its mathematical properties in detail. The unbiased regression estimator is as follows:

$$t_6 = \bar{y} + b_{yx}\left(\bar{X} - \bar{x}\right). \tag{2.2.16}$$

The mean square error of t_6 is

$$MSE(t_6) = \bar{Y}^2\theta C_y^2\left[1 - \rho^2{}_{yx}\right]. \tag{2.2.17}$$

Searls (1964) developed shrinkage estimator is as,

$$t_7 = \lambda\bar{y}, \tag{2.2.18}$$

where λ is a suitable constant. If $\lambda = 1$, the estimator t_7 becomes unbiased mean estimator. The bias and the *MSE* of t_7 are as follows,

$$Bias(t_7) = (\lambda - 1)\bar{Y}, \tag{2.2.19}$$

and

$$MSE(t_7) = \bar{Y}^2 \left[1 - \frac{1}{1+\theta C_y^2}\right]. \tag{2.2.20}$$

The *MSE* of t_7 is minimum if $\lambda = \frac{1}{1+\theta C_y^2}$.

Rao (1991) combined the idea of Cochran (1940) and Searl (1964), and presented shrinkage regression estimator as

$$t_8 = \alpha\bar{y} + b_{yx}(\bar{X} - \bar{x}), \tag{2.2.21}$$

where α and *b* are constants. The bias and *MSE* of t_8 are given by,

$$Bias(t_8) = (\alpha - 1)\bar{Y}, \tag{2.2.22}$$

and

$$MSE(t_8) = \bar{Y}^2 \left[1 - \frac{1}{1+\theta C_y^2\left(1-\rho^2{}_{yx}\right)}\right]. \tag{2.2.23}$$

The *MSE* of t_8 is minimum for $b_{yx} = \frac{\alpha\bar{Y}C_y\rho_{xy}}{\bar{X}C_x}$ and $\alpha = \frac{1}{1+\theta C_y^2\left(1-\rho_{yx}^2\right)}$. The estimator given by Rao (1991) was considered to be better estimate as compared to linear regression estimator.

Also a situation was found when exponential relation between study variable and auxiliary variable exists. For such a case, Bhal and Tuteja (1991) proposed exponential type ratio and product estimators as,

$$t_9 = \bar{y}\exp\left[\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right], \tag{2.2.24}$$

and

$$t_{10} = \bar{y}\exp\left[\frac{\bar{x} - \bar{X}}{\bar{x} + \bar{X}}\right]. \tag{2.2.25}$$

The bias of t_9 and t_{10} respectively are,

$$Bias(t_9) = \theta\bar{Y}C_x^2\left(\frac{3}{8} - \frac{1}{2}\rho_{yx}\frac{C_y}{C_x}\right), \tag{2.2.26}$$

and

$$Bias(t_{10}) = \theta\bar{Y}C_x^2\left(\rho_{yx}\frac{C_y}{2C_x} - \frac{3}{8}\right). \tag{2.2.27}$$

The *MSE*'s of the exponential ratio and product estimators are respectively given as,

$$MSE(t_9) = \bar{Y}^2\theta\left[C_y^2 + \frac{1}{4}C_x^2 - \rho_{yx}C_xC_y\right], \tag{2.2.28}$$

and

$$MSE(t_{10}) = \bar{Y}^2\theta\left[C_y^2 + \frac{1}{4}C_x^2 + \rho_{yx}C_xC_y\right]. \tag{2.2.29}$$

The exponential estimator developed by Bahl and Tuteja (1991) provided more precise results than mean per unit, ratio and product estimators if the relation between *y* and *x* is exponential.

Singh et al. (2008) presented a generalized exponential estimator motivated by Bahl and Tuteja (1991), the estimator is given as,

$$t_{11} = \bar{y}\left[\alpha\exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right) + (1-\alpha)\exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right)\right]. \tag{2.2.30}$$

The bias and the *MSE* of t_{11} are respectively presented as,

$$Bias(t_{11}) = \theta\bar{Y}C_x^2\frac{1}{2}\left(\frac{3}{4} - \rho_{yx}\frac{C_y}{C_x}\right), \tag{2.2.31}$$

and

$$MSE(t_{11}) = \bar{Y}^2\theta\left[C_y^2 + \frac{1}{4}C_x^2 - \rho_{yx}C_xC_y\right]. \tag{2.2.32}$$

Grover and Kaur (2011) followed Ray and Singh (1981) and Bahl and Tuteja (1981). They presented regression in exponential form as,

$$t_{12}=\left[d_1'\bar{y}+d_2'\left(\bar{X}-\bar{x}\right)\right]\exp\left[\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right], \tag{2.2.33}$$

where d_1' and d_2' are suitable constants. The bias and the *MSE* of t_{12} are respectively written as follows,

$$Bias\left(t_{12}\right)=\left[\left(d_1'-1\right)+\theta d_1'\frac{C_x}{2}\left(\frac{3}{4}C_x-\rho_{yx}C_y\right)\right]+d_2'\bar{X}\theta\frac{C_x^2}{2}, \tag{2.2.34}$$

and

$$MSE\left(t_{12}\right)=\frac{\theta\bar{Y}^2C_y^2\left(1-\rho_{yx}^2\right)}{1+\theta C_y^2\left(1-\rho_{yx}^2\right)}-\frac{\theta^2\bar{Y}^2C_x^2\left[4C_y^2\left(1-\rho_{yx}^2\right)+\frac{C_x^2}{4}\right]}{16\left[1+\theta C_y^2\left(1-\rho_{yx}^2\right)\right]}. \tag{2.2.35}$$

The proposed estimator of Grover and Kaur (2011) was considered to be more efficient estimator to estimate population mean of the study variable under numerical and mathematical comparisons with the estimators available in literature.

Yadav and Kadilar (2013) proposed a shrinkage exponential estimator is as,

$$t_{13}=w\bar{y}\exp\left[\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right]. \tag{2.2.36}$$

The estimator was presented by combining the idea of Searls (1964) shrinkage estimator and Bahl and Tuteja (1991) ratio-type exponential estimator. The expressions of the bias and the *MSE* of t_{13} are taken as,

$$Bias\left(t_{13}\right)=\bar{Y}\left(w-1\right)+\bar{Y}\theta\frac{C_x}{2}\left(\frac{3}{4}C_x-\rho_{yx}C_y\right), \tag{2.2.37}$$

and

$$\min MSE\left(t_{13}\right)=\bar{Y}^2\left[1-\frac{\left\{1+\theta\frac{1}{2}\left(\frac{3}{4}C_x^2-C_xC_y\rho_{yx}\right)\right\}^2}{1+\theta\left(C_y^2+C_x^2-2C_xC_y\rho_{yx}\right)}\right]. \tag{2.2.38}$$

Sharma et al. (2013), Grover and Kaur (2014), and many others worked on the estimation population mean of study variable when there is exponential relation is observed between *y* and *x*.

2.3 ESTIMATORS UNDER TWO-PHASE SAMPLING FOR NON-SENSITIVE STUDY VARIABLE

In the earlier discussion, different ratio, product and regression estimators along with their different modified and generalized forms have been presented. All such types of estimators require prior knowledge about the population parameter(s) of the auxiliary variable such as $\bar{X}$ whereas under some practical situations, the auxiliary information is not readily available. To obtain better estimators by using the relationship between auxiliary variable and the variable of interest, Neyman (1938) introduced two-phase sampling scheme to be useful. The usual ratio and regression estimators for two-phase sampling scheme under the situation when population mean $\bar{X}$ is unknown presented by Cochran (1977) are given as,

$$t_{14} = \frac{\bar{y}''}{\bar{x}''}\bar{x}', \tag{2.3.1}$$

and

$$t_{15} = \bar{y}'' + b_{yx}\left(\bar{x}' - \bar{x}''\right). \tag{2.3.2}$$

The bias of t_{14} is written as,

$$Bias\left(t_{14}\right) = \bar{Y}\theta_3\left(C_x^2 - \rho_{yx}C_yC_x\right). \tag{2.3.3}$$

The mean squared errors of t_{14} and t_{15} are given by,

$$MSE\left(t_{14}\right) = \bar{Y}^2\left[\theta_2C_y^2 + \theta_3\left(C_x^2 - 2\rho_{yx}C_yC_x\right)\right], \tag{2.3.4}$$

and

$$MSE\left(t_{15}\right) = \bar{Y}^2\theta_2C_y^2\left(1 - \rho_{yx}^2\right). \tag{2.3.5}$$

A generalized ratio estimator for two-phase sampling to estimate population mean of study variable was presented by Srivastava (1970) is given as

$$t_{16} = \bar{y}''\left(\frac{\bar{x}'}{\bar{x}''}\right)^{\alpha}, \tag{2.3.6}$$

where α is an unknown constant. The bias and the *MSE* of t_{16} are given as below,

$$Bias(t_{16}) = \bar{Y}\theta_3\alpha\left(\frac{(\alpha-1)}{2}C_x^2 - \rho_{yx}C_yC_x\right), \tag{2.3.7}$$

and

$$MSE(t_{16}) = \bar{Y}^2C_y^2\left(\theta_2 - \theta_3\rho_{yx}^2\right). \tag{2.3.8}$$

The generalized estimator t_{16} provides more precise results to estimate population mean than usual ratio estimator for two-phase sampling design.

The idea of chain ratio-type estimator in double sampling was introduced by Chand (1975) using two auxiliary variables under the situation when x_1 and x_2 have linear relationship with each other. He proposed the estimator as,

$$t_{17} = \frac{\bar{y}''}{\bar{x}_1''}\frac{\bar{x}_1'}{\bar{x}_2'}\bar{X}_2. \tag{2.3.9}$$

The bias and the *MSE* of t_{17} are written as,

$$Bias(t_{17}) = \bar{Y}\left[\theta_3\left(C_{x1}^2 - \rho_{yx1}C_yC_{x1} + \rho_{x1x2}C_{x1}C_{x2}\right) + \theta_2\left(C_{x2}^2 - \rho_{yx2}C_yC_{x2}\right)\right], \tag{2.3.10}$$

and

$$MSE(t_{17}) = \bar{Y}^2\left[\theta_2C_y^2 + \theta_3\left(C_{x_1}^2 - 2\rho_{yx_1}C_yC_{x_1}\right) + \theta_1\left(C_{x2}^2 - 2\rho_{yx_2}C_yC_{x2}\right)\right]. \tag{2.3.11}$$

The proposed estimator presented by Chand (1975) provided more precise results than usual ratio estimator for two-phase simple random sampling. Kiregyera (1984) introduced regression type estimator using two auxiliary variables. He considered the relationship between study (y) and auxiliary variable (x_1) was linear but not passing through the origin, on the other hand the relationship between x_1 and x_2 are linear and passes through the origin. The proposed estimator is given as,

$$t_{18} = \bar{y}'' + b_{yx_1}\left[\left(\bar{x}_1' - \bar{x}_1''\right) - b_{x_1x_2}\left(\bar{x}_2' - \bar{x}_2''\right)\right]. \tag{2.3.12}$$

The MSE of t_{18} up to first order of approximation is as,

$$MSE\left(t_{18}\right)=\bar{Y}^{2}C_{y}^{2}\left[\theta_{2}-\theta_{3}\rho_{yx_{1}}^{2}+\theta_{1}\rho_{yx_{1}}\rho_{x_{1}x_{2}}\left(\rho_{yx_{1}}\rho_{x_{1}x_{2}}-\rho_{yx_{2}}\right)\right]. \quad (2.3.13)$$

The proposed estimator's performance was examined using super population and found to be better precise estimate than usual mean and regression estimators. Sahoo et al. (1994) introduced two alternative estimators under two-phase sampling scheme when the population mean on the second auxiliary information is known. The estimators are written as below,

$$t_{19}=t_{14}+b_{yx2}\left(\bar{x}_{2}'-\bar{x}_{2}''\right), \quad (2.3.14)$$

and

$$t_{20}=t_{15}+b_{yx2}\left(\bar{x}_{2}'-\bar{x}_{2}''\right), \quad (2.3.15)$$

where a is a suitable constant and b_{yx2} is a sample regression coefficient if y on x_2. The MSE of t_{19} and t_{20} are given as,

$$MSE\left(t_{19}\right)=\theta_{2}S_{y}^{2}+2\frac{\bar{Y}}{\bar{X}_{1}}\theta_{3}\rho_{yx1}S_{y}S_{x1}-\frac{\bar{Y}^{2}}{\bar{X}^{2}}\theta_{3}S_{x1}^{2}+\theta_{1}\rho^{2}{}_{yx2}S_{y}^{2}, \quad (2.3.16)$$

and

$$MSE\left(t_{20}\right)=\left[\theta_{2}\left(1-\rho_{yx1}^{2}\right)+\theta_{1}\rho_{yx1}^{2}\right]S_{y}^{2}+\theta_{1}\rho_{yx2}^{2}S_{y}^{2}. \quad (2.3.17)$$

The estimator t_{20} works efficiently than t_{15} and Chand (1975) chain ratio estimator if y and x_1 have linear relation and the line passes through the origin, otherwise t_{20} is suitable to use and performs better than Cochran (1977) regression estimator and Kiregyera (1984) estimator.

Singh et al. (1994) suggested ratio and product class of estimators to estimate means of finite populations using two auxiliary variables. Singh and Upadhaya (1995), Singh and Singh (1991), Tracy and Singh (1997, 1999), Singh and Gangele (1999),Singh and Espejo (2000), Khoshnevisan et al. (2002), Singh and Espejo (2003), Roy (2003), Singh et al. (2006) and Sammudin and Hanif (2007) presented some modified estimators for ratio, product and regression under two-phase sampling scheme.

Additionally, Singh and Vishwakarma (2007) modified the ratio and product exponential estimators of Bhal and Tuteja (1991) for two-phase sampling for the case when the population mean of auxiliary variable was unknown. The modified estimators under the situation when population mean of *y* and *x* are unknown are respectively given by,

$$t_{21} = \bar{y}'' \exp\left(\frac{\bar{x}' - \bar{x}''}{\bar{x}' + \bar{x}''}\right), \tag{2.3.18}$$

and

$$t_{22} = \bar{y}'' \exp\left(\frac{\bar{x}'' - \bar{x}'}{\bar{x}'' + \bar{x}'}\right). \tag{2.3.19}$$

The *MSE*'s of estimator t_{21} and t_{22} are written by,

$$MSE\left(t_{21}\right) = \bar{Y}^2 C_y^2 \left[\theta_2 - \theta_3 \frac{C_x}{C_y}\left(\frac{C_x}{C_y} - 2\rho_{yx}\right)\right], \tag{2.3.20}$$

and

$$MSE\left(t_{22}\right) = \bar{Y}^2 C_y^2 \left[\theta_2 - \theta_3 \frac{C_x}{C_y}\left(\frac{C_x}{C_y} + 2\rho_{yx}\right)\right]. \tag{2.3.21}$$

Singh and Choudhary (2012) proposed chain ratio and product type estimators to estimate population mean of study variable in two-phase sampling when information on main auxiliary variable is unknown but the information on second auxiliary variable is available. The estimators are respectively given as,

$$t_{23} = \bar{y}'' \exp\left(\frac{\bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2'} - \bar{x}_1''}{\bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2'} + \bar{x}''}\right), \tag{2.3.22}$$

and

$$t_{24} = \bar{y}'' \exp\left(\frac{\bar{x}_1'' - \bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2'}}{\bar{x}_1'' + \bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2'}}\right). \tag{2.3.23}$$

The bias of estimator's t_{23} and t_{24} are presented as,

$$Bias\left(t_{23}\right) = \bar{Y}\left[\frac{3}{8}\left(\theta_3 C_{x1}^2 + \theta_1 C_{x2}^2\right) + \frac{1}{2}\left(\theta_3 C_{x1}^2 C_{yx1} + \theta_1 C_{yx2} C_{x2}^2\right)\right], \tag{2.3.24}$$

and

$$Bias(t_{24}) = \bar{Y}\left[-\frac{1}{8}\left(\theta_3 C_{x1}^2 + \theta_1 C_{x2}^2\right) + \frac{1}{2}\left(\theta_3 C_{x1}^2 C_{yx1} + \theta_1 C_{yx2} C_{x2}^2\right)\right]. \quad (2.3.25)$$

The *MSE* of estimator's t_{23} and t_{24} up to first order of approximation are as,

$$MSE(t_{23}) = \bar{Y}^2\left[\theta_2 C_y^2 + \frac{1}{4}\left(\theta_3 C_{x1}^2 + \theta_1 C_{x2}^2\right) - \theta_3 C_{x1}^2 C_{yx1} - \theta_1 C_{yx2} C_{x2}^2\right], \quad (2.3.26)$$

and

$$MSE(t_{24}) = \bar{Y}^2\left[\theta_2 C_y^2 + \frac{1}{4}\left(\theta_3 C_{x1}^2 + \theta_1 C_{x2}^2\right) + \theta_3 C_{x1}^2 C_{yx1} + \theta_1 C_{yx2} C_{x2}^2\right]. \quad (2.3.27)$$

The above discussed estimators presented by Singh and Choudhary (2011) were examined to be efficient estimators than ratio, product and sample mean per unit estimator.

Khare et al. (2013) presented a generalized regression type estimator using two auxiliary variables and their proposed estimator is the form of a generalized ratio in regression estimator combining Srivastava (1970) generalized estimator in usual regression estimator and is given as,

$$t_{25} = \bar{y}'' + b_{yx_1}\left[\bar{x}_1'\left(\frac{\bar{X}_2}{\bar{x}_2'}\right)^{\alpha} - \bar{x}_1''\right], \quad (2.3.28)$$

where α is an unknown constant. The *MSE* of the above estimator is,

$$MSE(t_{25}) = \bar{Y}^2 C_y^2\left[\theta_2 - \theta_3\rho_{yx_1}^2 + \theta_1\rho_{yx_2}^2\right]. \quad (2.3.29)$$

Khare et al. (2013) provided some mathematical and numerical comparisons and concluded that the proposed estimator performs better than usual ratio, regression, Srivastava (1970) generalized estimator and Kiregyera (1984) estimator.

Khan (2016) proposed a ratio-type exponential estimator to estimate population mean of the study variable

$$t_{26} = \bar{y}''\left(\frac{\bar{x}' - \bar{x}_1''}{\bar{x}' + \bar{x}_1''}\right)^{k_1} + k_2\left[\bar{x}_1' \exp\left(\frac{\bar{X}_2 - \bar{x}_2'}{\bar{X}_2 + \bar{x}_2'}\right) - \bar{x}_1''\right], \tag{2.3.30}$$

where k_1 and k_2 are unknown constants. The mean squared error of t_{21} estimator is as,

$$MSE(t_{26}) = \bar{Y}^2 C_y^2\left[\theta_2 - \theta_3\rho_{yx_1}^2 - \frac{\theta_1\left(2\rho_{yx_1}C_{x_1} - \rho_{yx_2}C_{x_2}\right)^2}{\left(4C_{x_1}\left(C_{x_1} - \rho_{yx_2}C_{x_2}\right) + C^2{}_{x_2}\right)}\right]. \tag{2.3.31}$$

The estimator presented by Khan (2016) was efficient than some existing estimators and suggested to be useful in practice.

2.4 SOME EXISTING RANDOMIZED RESPONSE MODELS

The above argument on two-phase sampling estimation can be done when the observed variables have full response or non-response is negligible. The direct method (*DM*) to obtain information is not easy for sensitive surveys. So, to deal with this problem one can use randomized response (*RR*) technique pioneered by Warner (1965) in which a randomized device like a deck of cards, spinner, etc. is required. Consider a population in which each individual either belongs to sensitive group *A* or non-sensitive group *B*. The interest of the researcher is to estimate population proportion of the sensitive group *A*. A sample of size *n* is drawn under *SRSWR* from the population. Each interviewer is equipped with a randomization device before the interview. The probability of appearing in group *A* is *p* and the probability of appearing in group *B* is (1-*p*). So, the interviewee is asked to report whether or not the true group appears using randomization device and is unnoticed by the interviewer. The interviewee has to report if “yes” or “no” according to his/her true group *A* or *B*. So the model is as

$$\hat{\pi}_1 = \frac{\hat{\tau} + (p-1)}{2p-1}, \tag{2.4.1}$$

where $\hat{\tau}$ is the observed proportion of ‘*yes*’ response in the sample. And the sample variance is as

$$Var(\hat{\pi}_1) = \frac{\pi(1-\pi)}{n} + \frac{p(1-p)}{n(2p-1)^2}. \tag{2.4.2}$$

Following Warner (1965), a theoretical framework upon sensitive surveys was introduced by Greenberg et al. (1969). They adapted Simmon's method of unrelated questions and used in sensitive surveys to obtain responses for two samples as an extension of Warner's work. In the proposed unrelated question *RR* methods, two questions are asked from respondent in which one question is sensitive in nature and the other is inoffensive in nature. The probabilities of responding sensitive question are p_1 and p_2 and the probabilities to respond the unrelated question are $\left(1-p_1\right)$ and $\left(1-p_2\right)$. The unbiased estimate of the probability of the sensitive attribute is as,

$$\hat{\pi}_2 = \frac{\hat{\tau}_1\left(1-p_2\right)-\hat{\tau}_2\left(1-p_1\right)}{p_1-p_2}, \tag{2.4.3}$$

where $\hat{\tau}_1$ is the sample 1 observed proportions of 'yes' responses and $\hat{\tau}_2$ is the sample 2 observed proportion of 'yes' responses. The sample variance of the estimate is as,

$$Var\left(\hat{\pi}_2\right)=\frac{\left[\tau_1\left(1-\tau_1\right)\left(1-p_2\right)^2/n_1\right]+\left[\tau_2\left(1-\tau_2\right)\left(1-p_2\right)^2/n_2\right]}{\left(p_1-p_2\right)^2}. \tag{2.4.4}$$

Greenberg et al. (1969) model provided more efficient estimator than Warner (1965) model and also the untruthful response rate was observed to be reduced.

Warner (1971) introduced a general linear randomized response model for multivariate mixes of randomized and non-randomized response using continuous and discrete random variables. The observed reported responses is

$$Z_{1_i} = T_i y_i, \qquad i=1,2,\ldots,n. \tag{2.4.5}$$

where T_i is $r\times p$ random matrix of known distribution and y_i represents the observation on *Y*. It assumed that the above product matrix is accurate but the actual values of T_i and y_i remains hidden to the interviewer. Warner (1971) also presented the applications of the proposed model.

Greenberg et al. (1971) developed the *RR* model to reduce the response bias under the situation where quantitative sensitive data is available. They proposed a quantitative randomized response model to estimate induced abortion rates in urban area a randomization device is used on the data of North Carolina Abortion Study. The observed responses for the two independent non-overlapping samples n_1 and n_2 are as,

$$Z_2 = \frac{(1-p_2)\bar{T}_1 - (1-p_1)\bar{T}_2}{p_1 - p_2}, \text{ and } Z_3 = \frac{p_2\bar{T}_1 - p_1\bar{T}_2}{p_2 - p_1}. \tag{2.4.6}$$

The variance can be written as,

$$Var\left(Z_{(i)}\right) = \frac{1}{n_{(i)}p_{(i)}}\left(S_y^2 + p_{(i)}\left(S_s^2 - S_y^2\right) + p_{(i)}\left(1 - p_{(i)}\right)\left(\bar{S} - \bar{Y}\right)^2\right). \; i=2, 3. \tag{2.4.7}$$

The data from North Carolina Abortion Survey was used to illustrate the proposed procedure. Then Pollock and Bek (1976) extended the work of Greenberg et al. (1969) for unrelated question model. They proposed two models for unrelated question for *RR* methods which involves addition and multiplication of random number generated from the known distribution to the sensitive variable of interest. The reported response for additive model is as,

$$Z_4 = Y + S. \tag{2.4.8}$$

The mean and variance of Z_4 are,

$$E_M\left(Z_4\right) = \bar{Y} + \bar{S}, \tag{2.4.9}$$

and

$$Var\left(Z_4\right) = S_y^2 + S_s^2. \tag{2.4.10}$$

The multiplicative model presented as,

$$Z_5 = YR, \tag{2.4.11}$$

where R is another scrambling variable generated from some known distribution. The mean and the variance for multiplicative model are,

$$E_M\left(Z_5\right) = \overline{YR}, \tag{2.4.12}$$

and

$$Var\left(Z_5\right) = S_R^2\left(S_y^2 + \bar{Y}^2\right) + S_y^2\bar{R}^2. \tag{2.4.13}$$

The moments and the parametric estimations are also presented for each model discussed and compared the variances with Greenberg et al. (1969)

model's' variance. They showed that the additive and multiplicative model were more efficient.

Adapting Pollock and Bek (1976) additive model, Himmelfarb and Edgell (1980) introduced general additive constant models for quantitative sensitive questions. The general case for the additive model is given as,

$$Z_{6(i)} = Y + S_{(i)}, \qquad i = 1, 2, \ldots, c. \tag{2.4.14}$$

The mean and the variance for the above model are as,

$$E_M\left(Z_{6(i)}\right) = \bar{Y} + \sum p_i S_{(i)}, \tag{2.4.15}$$

and

$$Var\left(Z_{6_{(i)}}\right) = S_y^2 + \sum p_{(i)} S_{(i)}^2 + \left(\sum p_{(i)} S_{(i)}\right)^2, \tag{2.4.16}$$

where p_i is the probability of occurrence of *S* for *c* different constants. The special cases for generalized additive model were assumed by substituting different values for *S* i.e., *2S* or $S(1/4)$. Himmelfarb and Edgel general additive model and the special cases were comparatively efficient than Greenberg et al. (1969) unrelated question method.

After Himmelfarb and Edgel (1980), Eichhorn and Hayre (1983) revisited Pollock and Bek (1971) multiplicative *SRR* technique for obtaining responses on the sensitive variable when responses are quantitative in nature. The multiplicative *SRR* technique involves a random device number having some known distribution and the respondents answer is multiplied by the random number and the resulted product is provided to the interviewer, who is unaware of the random number chosen by the respondent. So the scrambled response is received by the interviewer. The *SRR* model is presented as,

$$Z_7 = YR. \tag{2.4.17}$$

The mean and the variance of *SRR* model are written as,

$$E_M\left(Z_7\right) = \overline{YR}, \text{ and } Var\left(Z_7\right) = S_R^2\left(S_y^2 + \bar{Y}^2\right) + S_y^2 \bar{R}^2. \tag{2.4.18}$$

They discussed different ways to generate scrambling numbers. The *SRR* model is considered to be the special case of Warner (1971) general linear randomized response model. Eichhorn and Hayre (1983) discussed some theoretical properties of the *SRR* model comparison with unrelated question method presented by Greenberg et al. (1971). The observed model Z_7 was found to be superior than Greenberg et al. (1971) method. They also presented an idea of using two scrambled response as $Z = RY + S$, a combination of multiplicative and additive approach for extreme sensitive surveys.

Singh et al. (1996) presented a general linear regression model $Y = X\beta + \varepsilon$ assuming the dependent variable is observed as a scrambling response via Eichhorn and Hayre's (1983) *SRR* approach. Simulation study was performed to compare the estimates of study variable obtained from scrambling variables with conventional or direct question surveys. The *SRR* method was then adapted by Strachan et al. (1998) to evaluate the performance of the likelihood based estimator for linear regression model with scrambled responses. They used two estimation methods i.e., maximum likelihood estimation and Bayesian estimation.

Gupta et al. (2002) introduced an Optional randomized response (*ORR*) model an alternative procedure of *RR* technique for sensitive surveys. In this intended technique an option is provided to the respondent to answer either true or scrambled response and the interviewer is unaware of the response type chosen by the interviewee. So the reported response for *ORR* model is given as,

$$Z_8 = YS^k, \tag{2.4.19}$$

where k is a Bernoulli variate with mean $E(k) = W$ and W is the sensitivity level of the question. If $k = 0$, the model provides true response and if $k = 1$, Z_8 gives scrambled response as presented by Eichhorn and Hyare (1983). The unbiased estimator for *ORR* model with its variance are written as,

$$t_{27} = \frac{1}{n}\sum_{i=1}^{n} Z_{8i} \text{ , and } Var(t_{27}) = \frac{1}{n}\left[S_y^2 + WS_s^2\left(S_y^2 + \bar{Y}^2\right)\right]. \tag{2.4.20}$$

The purpose of the proposed model was to estimate the population mean of the sensitive variable as well as estimate the sensitivity level.

Gupta and Shabbir (2004) presented an *ORR* model. They used double sampling approach in their proposed model, the reposted response is taken as,

$$Z_{9(i)} = Y(1 - W) + YS_{(i)}W, \qquad i = 1,2. \tag{2.4.21}$$

The mean derived for *ORR* model by Gupta et al. (2004) is written as,

$$E_M\left(Z_{9(i)}\right)=\left[\bar{S}_{(i)}W+\left(1-W\right)\right]\bar{Y}\,. \tag{2.4.22}$$

The unbiased estimator for which the *ORR* model of Gupta and Shabbir (2004) with its variance are,

$$t_{28}=\frac{\bar{Z}_{9(1)}\left(\bar{S}_{(2)}-1\right)-\bar{Z}_{9(2)}\left(\bar{S}_{(1)}-1\right)}{\bar{S}_{(2)}-\bar{S}_{(1)}}, \tag{2.4.23}$$

and

$$Var\left(t_{28}\right)=\frac{\left[S_{z_1}\left(\bar{S}_{(2)}-1\right)+S_{z_2}\left(\bar{S}_{(1)}-1\right)\right]^2}{n\left(\bar{S}_{(2)}-\bar{S}_{(1)}\right)^2}. \tag{2.4.24}$$

The proposed *ORR* model of Gupta and Shabbir (2004) estimates both $\bar{Y}$ as well as *W*, unlike forced *SRR* model. They discussed the purpose of the proposed model and a numerical example was presented to illustrate simulation results. They justified that the estimator based *ORR* model is efficient than non-optional *RR* technique.

Bar-lev et al. (2004) introduced a generalized version of Eichhorn and Hayre (1983) multiplicative *SRR* model including design parameter. The reported response is given as,

$$Z_{10}=\begin{cases} Y & \text{with probability } p \\ YS & \text{with probability } \left(1-p\right) \end{cases}, \tag{2.4.25}$$

where $p\epsilon\left[0,1\right]$, controlled by the researcher and used for randomizing the responses. If $p=0$, the above model becomes Eichhorn and Hayre (1983) *SRR* model and if the choice is made to $p=1$, the model reduces to direct response interview. The proposed generalized model has minimum variance than Eichhorn and Hayre (1983) multiplicative model.

The two-stage quantitative randomized response model we proposed by Ryu et al. (2005).The proposed model was presented for *SRSWR* as well as stratified random sampling design. The reported response is observed as,

$$Z_{11}=\alpha Y+\left(1-\alpha\right)\left(\beta Y+\left(1-\beta\right)YS\right), \tag{2.4.26}$$

where α and β are Bernoulli random variables. The expected value of the observed response is as

$$E_M\left(Z_{11}\right)=\bar{Y}. \tag{2.4.27}$$

The simulation study was accomplished to observe the efficiency of the proposed model. In terms of variance, Ryu et al. (2005) model was noted to be more efficient as compared to Greenberg et al. (1971) model.

Guerriero and Sandri (2007) made comparisons on some *RR* models based on their efficiency and privacy protection criterion. They found that some *RR* methods are perfectly equitant to Warner's model and some are better than Warner's model. Saha (2007) introduced an independent *RR* device useful for both dichotomous as well as quantitative responses from the respondents. The suggested *RR* model was the combination of multiplicative and additive models is taken as,

$$Z_{12}=R\left(Y+S\right), \tag{2.4.28}$$

where S and R are the two scrambled randomized response variables with known distribution. The mean and the variance of Z_{12} are as follows,

$$E_M\left(Z_{12}\right)=\bar{R}\left(\bar{Y}+\bar{S}\right), \tag{2.4.29}$$

and

$$Var\left(Z_{12}\right)=\left(\bar{Y}^2+S_y^2\right)S_R^2+\left(S_R^2S_S^2+S_R^2\bar{S}^2+S_S^2\bar{R}^2\right)+2\overline{YS}S_R^2. \tag{2.4.30}$$

This model Z_{12} was better than Warner (1965) model and Eichhorn and Hayre (1983) model. Also, Saha (2007) suggested that the model Z_{12} can be implemented in stratified unequal probability sampling. Saha (2008) used Eichhorn and Hayre (1983) *SRR* model and used in unequal probability sampling design. Also, a numerical example was presented to examine the performance of the proposed procedure under two conventional sampling designs.

A useful *RR* model for quantitative sensitive surveys was introduced by Gjestvan and Singh (2009) to estimate population mean of sensitive variable. The observed response is taken as,

$$Z_{13} = \begin{cases} Y + aS \quad with\,probabilty\;\; p = \dfrac{\beta}{\alpha+\beta} \\ Y - \beta S \quad with\,probability\left(1-p\right) = \dfrac{\alpha}{\alpha+\beta} \end{cases}, \tag{2.4.31}$$

The unbiased estimator using the proposed *RR* model is written as,

$$t_{29} = \frac{1}{n}\sum_{i=1}^{n} z_{13i}\,, \tag{2.4.32}$$

The variance becomes,

$$Var\left(t_{29}\right) = \frac{1}{n}\left[S_y^2 + \alpha\beta\left(S_s^2 + \bar{S}^2\right)\right]. \tag{2.4.33}$$

Gjestvang and Singh (2009) used the data on the average *GPA* of students at St. Cloud State University and generated random numbers for *S*. They applied the proposed model to examine the efficiency. From the theoretical and empirical results of students scrambled responses, it was concluded that the suggested model is safely and securely useful.

The efficiency of different strategies on quantitative characteristics of sensitive variables using *SRR* technique was presented by Yan et al. (2009). They focused on three different *RR* models i.e., Greenberg et al. (1971), Eichhorn and Hayre (1983) and Gupta et al. (2002) for the comparisons of efficiency and protection degree to the respondents. The protection degree is defined as,

$$\Delta = E\left(Z - Y\right)^2, \tag{2.4.34}$$

where *Z* is a randomized response and *Y* is the respondent's true value of the quantitative sensitive character. It was noticed by comparison that at the same level of protection degree, the model of Eichhorn and Hayre (1983) has statistically equivalent strategy as that of Gupta et al. 's (2002) model, but Greenberg et al. (1971) performs less efficient as compared to Eichhorn and Hayre (1983) and Gupta et al. (2002) *RR* strategies. Chaudhari et al. (2009) discussed six different *RR* models presented in literature on efficiency and privacy protection in *RR*. They made extension on these models to general sample selection methods. They presented privacy measures of jeopardy and related quantities for the six *RR* models.

Bouza et al. (2010) reviewed different *RR* models for qualitative sensitive surveys and applied Monte Carlo experiment to make comparisons among *RR* models under study but that would not bring out any conclusion. Motivated from Saha (2007), three different *SRR* models were presented by Diana and Perri (2010). Their first generalized model when the interviewee is asked to respond the sensitive question with the probability p is as follows,

$$Z_{14} = Yp + R(Y+S)(1-p). \tag{2.4.35}$$

If $p=0$, then the model Z_{14} becomes Saha (2007) model and if $p=1$, the true responses will be collected from *DM*. The mean of the model Z_{14} becomes,

$$E(Z_{14}) = \bar{Y}p + \bar{R}(\bar{Y}+\bar{S})(1-p). \tag{2.4.36}$$

The estimator using Z_{14} is taken as,

$$t_{30} = \frac{\bar{z}_{14} - \bar{R}\bar{S}(1-p)}{p+(1-p)\bar{R}}. \tag{2.4.37}$$

And the variance of the estimator of t_{30} is as,

$$Var(t_{30}) = \frac{S^2_{Z_{14}}}{n\left[p+(1-p)\bar{R}\right]^2}. \tag{2.4.38}$$

The second *SRR* model by Diana and Perri (2010) was proposed by combining additive and multiplicative approaches. Their model observed is as,

$$Z_{15} = R\left[\alpha S + (1-\alpha)Y\right], \tag{2.4.39}$$

where α is a suitable constant. If $\alpha = 0$, then the model Z_{15} becomes Eichhorn and Hayre (1983) multiplicative model The mean of the second *SRR* model is as,

$$E(Z_{15}) = +\alpha\bar{R}\bar{S} + (1-\alpha)\overline{YR}. \tag{2.4.40}$$

The mean estimator using Z_{15} is given as,

$$t_{31} = \frac{\bar{z}_{15} - \alpha \bar{R}\bar{S}}{(1-\alpha)\bar{R}}. \quad (2.4.41)$$

And the variance of the estimator of t_{31} is as,

$$Var(t_{31}) = \frac{S_{Z_{15}}^2}{n\left[(1-\alpha)\bar{R}\right]^2}. \quad (2.4.42)$$

Another generalized *SRR* model by Diana and Perri (2010) is

$$Z_{16} = \tau(Y+S) + (1-\tau)YR. \quad (2.4.43)$$

where τ is a suitable constant. For $\tau = 0$, the model Z_{16} reduces to Eichhorn and Hayre (1983) multiplicative model and if $\tau = 1$ the model generates Pollock and Bek’s (1971) additive model. The mean of the third generalized *SRR* model is as,

$$E(Z_{16}) = \tau(\bar{Y} + \bar{S}) + (1-\tau)\overline{YR}. \quad (2.4.44)$$

The mean estimator and its variance under the model Z_{16} are respectively as,

$$t_{32} = \frac{\bar{z}_{16} - \tau\bar{S}}{\tau + (1-\tau)\bar{R}} \quad (2.4.45)$$

and

$$Var(t_{32}) = \frac{S_{Z_{16}}^2}{n\left[\tau + (1-\tau)\bar{R}\right]^2} \quad (2.4.46)$$

They discussed the properties of each of the proposed *SRR* models and also the privacy protection measures were provided for each model.

Following *ORR* approach of Gupta et al. (2002), an extension was made by Gupta et al. (2010) using split sample approach. They developed a model on the assumption that population mean of study variable (y) and sensitivity level say W can be estimated. The two sub-sample reported responses are given as,

$$Z_{17(i)} = \begin{cases} Y & with\ probabilty\ T + (1-T)(1-W) \\ Y + S_{(i)} & with\ probability\ (1-T)W \end{cases}, \ i=1,2. \tag{2.4.47}$$

where T is the proportion of respondents who were asked to provide truthful responses. The means of Z is as,

$$E_M\left(Z_{17(i)}\right) = \bar{Y} + S^2_{s(i)}W(1-T). \tag{2.4.48}$$

An unbiased estimator for mean given by Gupta et al. (2010) using $Z_{17(i)}$ is given as,

$$t_{33} = \frac{\bar{S}_{(1)}\bar{z}_{17(2)} - \bar{S}_{(2)}\bar{z}_{17(1)}}{\bar{S}_{(1)} + \bar{S}_{(2)}}, \text{ and } \hat{W} = \frac{\bar{z}_{17(2)} - \bar{z}_{17(1)}}{\left(\bar{S}_{(1)} - \bar{S}_{(2)}\right)\left(1-T\right)}. \tag{2.4.49}$$

The variance obtained for t_{33} estimator is presented as,

$$Var\left(t_{33}\right) = \frac{1}{\left(\bar{S}_{(1)} + \bar{S}_{(2)}\right)^2}\left[\bar{S}^2_{(2)}\frac{S^2_{z_1}}{n_1} + \bar{S}^2_{(1)}\frac{S^2_{z2}}{n_2}\right]. \tag{2.4.50}$$

The simulation study was performed and noticed that two-stage *ORR* model is efficient as compared to one-stage *RR* model. Gupta et al. (2010) also presented the advantages and disadvantages of multiplicative and additive *SRR* models. Diana and Perri (2011) introduced some *SRR* models to collect information by securing respondents privacy. They presented models for sensitive study variable as well as sensitive auxiliary variable. The privacy protection measure for the class of estimators were also provided.

Chang and Kuo (2012) considered the problem of estimating the proportion of a sensitive variable. They applied Wilson (1927) score method to construct the confidence intervals and the point estimates for the proposed estimator. Chen and Singh (2012) developed a scrambled model to estimate multinomial proportions of sensitive variable. The proposed model is as

$$Z_{18} = \begin{cases} S_1 & with\ probability \quad \pi_1 \\ S_2 & with\ probability \quad \pi_2 \\ S_3 & with\ probability \quad \pi_3 \end{cases}, \tag{2.4.51}$$

where S_1, S_2 and S_3 are random numbers with known distribution and π_1, π_2 and π_3 are the proportions of person belongs to sensitive group 1, 2 and 3. The mean an the variance of Chan and Singh presented *SRR* model are written as,

$$E_M\left(Z_{19}\right)=\pi_1\bar{S}_1+\pi_2\bar{S}_2+\left(1-\pi_1-\pi_2\right)\bar{S}_3, \tag{2.4.52}$$

and

$$Var\left(Z_{19}\right)=\left(\sigma_{S_1}^2-\sigma_{S_2}^2\right)\pi_1+\left(\sigma_{S_2}^2-\sigma_{S_3}^2\right)\pi_2+\pi_1\left(1-\pi_1\right)\left(\bar{S}_1-\bar{S}_3\right)^2$$
$$+\pi_2\left(1-\pi_2\right)\left(\bar{S}_2-\bar{S}_3\right)^2-2\left(\bar{S}_1-\bar{S}_2\right)\left(\bar{S}_2-\bar{S}_3\right)\pi_1\pi_2+\sigma_{S_3}^2. \tag{2.4.53}$$

Hussain (2012) proposed a new *RR* model to avoid the response bias in sensitive survey responses. The proposed model is as,

$$Z_{20}=lY+\frac{1}{l^l}\bar{S}, \tag{2.4.54}$$

where *S* is unrelated variable and *l* is a fixed constant. This constant *l* provides greater flexibility to the proposed model. The mean of the presented model is as,

$$E_M\left(Z_{20}\right)=l\bar{Y}+\frac{1}{l^l}\bar{S}. \tag{2.4.55}$$

The unbiased estimator follow as previously mentioned can be modified as,

$$t_{34}=l^{-1}\left[\bar{z}_{20}+\frac{1}{l^l}\bar{S}\right], \tag{2.4.56}$$

and the variance of the above estimator is given as,

$$Var\left(t_{34}\right)=\frac{1}{n}\left[S_y^2+\frac{1}{nl^{2l+3}}S_s^2\right]. \tag{2.4.57}$$

Unconditionally, the proposed model performed more efficiently as compared to Hussain and Shabbir (2007) *RR* model.

Hussain and Khan (2013) presented two models using additive scrambling technique to estimate the population mean of sensitive variable. For the models two different scrambling variables are assumed from some known distribution. The two reported response for the *ith* respondents given by Hussain and Khan (2013) are written as,

$$Z_{21(i)} = Y_{(i)} + S_{1(i)}, \text{ and } Z_{22(i)} = Y_{(i)} + S_{2(i)}, \text{ where } i = 1, 2, ..n. \qquad (2.4.58)$$

The means of the above models are,

$$E_M\left(Z_{21(i)}\right) = \bar{Z}_{21} = \bar{Y}, \text{ and } E_M\left(Z_{22(i)}\right) = \bar{Z}_{22} = \bar{Y}. \qquad (2.4.59)$$

From the models (2.3.58), Hussain and Khan (2013) presented an estimator is given as,

$$t_{35} = \lambda\bar{z}_{21} + (1-\lambda)\bar{z}_{22}. \qquad (0 < \lambda \le 1) \qquad (2.4.60)$$

The variance of the above estimator t_{35} is as,

$$Var(t_{35}) = \frac{1}{2n}\left(S_y^2 + S_s^2\right) + \frac{1}{2}S_{\bar{Z}_{21}\bar{Z}_{22}}. \qquad (2.4.61)$$

Also, the degree of confidentiality for each of the proposed models was derived mathematically using Zaizai et al. (2009) protection measure. They established that scrambling the responses increases the variance of the estimates.

Dihidar and Chowdhary (2013) focused on the problem of estimating population mean of variables such as expenditure on alcohol, abortion, amount of dowry, etc. Instead of using *SRSWR* scheme, they discussed unbiased estimator for varying probabilities sampling scheme based on Gjestvang and Singh (2009) *RR* model presented in (2.3.31). Their proposed estimator under Gjestvang and Singh (2009) *RR* model is given by,

$$t_{36} = \sum_{i\epsilon s} m_i b_{si}, \qquad (2.4.62)$$

where b_{si} are independent values of y_i values and $\sum_{s\ni i} p(s)b_{si} = 1$. Also, $E(m_i) = Y_i$. The variance is written as,

$$V(t_{36}) = \sum_{i\epsilon s} y_i^2 C_{si} + \sum_{i\epsilon s}\sum_{j\epsilon s, i\ne j} y_i y_j C_{sij} + E\left[\sum_{i\epsilon s} b_i^2 V_R(m_i)\right]. \qquad (2.4.63)$$

The proposed estimator was compared to Gjestvang and Singh (2009) estimator and found to be efficient in terms of variance.

Hussain et al. (2014) suggested additive and subtractive scrambled *RR* model using the *RR* model presented by Mehta et al. (2012). They implemented split approach and double sample approach. They derived unbiased mean estimator, variance and the sensitivity level based on the proposed model. The comparison was made among the proposed procedure, Mehta et al. (2012), Huang (2010) and Gupta et al. (2010) models. They obtained improved estimation of mean of the study variable using the proposed model.

Tarray and Singh (2015) presented improved versions of three different models such as Hummelfarb and Edgel (1980), Gjestvang and Singh (2009) and Singh (2010) additive models. The improved proposed model is as,

$$Z_{30(i)} = Y_{(i)} + S_{\omega}, \quad \text{where } i = 1, 2, ..n. \tag{2.4.64}$$

where $S_{\omega} = \omega\mu_s S^{*2} + (1-\omega)S$, $S^{*} = \dfrac{S-\mu_s}{\sigma_s}$ and ω is a suitable chosen scalar. The mean and the variance of above model are given by,

$$E_M\left(Z_{30(i)}\right) = \bar{Y} + \bar{S}, \tag{2.4.65}$$

and

$$Var\left(Z_{30(i)}\right) = \frac{1}{n}\left[S_y^2 + S_s^2\left\{\frac{\Upsilon(S)}{\Upsilon(S) + (C_s - \nu_1)^2}\right\}\right]. \tag{2.4.66}$$

Singh and Gorey (2016) suggested a generalized *RR* models which uses known values of mean and variance of the scrambling variable. The proposed model focused to collect true response to a sensitive question by protecting the respondent’s privacy. The proposed model is given by,

$$Z_{31} = \begin{cases} Y + \alpha S\left(\dfrac{2S_s\bar{S}}{S_s^2 + \bar{S}^2}\right)^{1/2} & \text{with probability } p = \dfrac{\beta}{\alpha+\beta} \\ Y - \beta S\left(\dfrac{2S_s\bar{S}}{S_s^2 + \bar{S}^2}\right)^{1/2} & \text{with probability } (1-p) = \dfrac{\alpha}{\alpha+\beta} \end{cases}, \tag{2.4.67}$$

where α and β are two known positive real numbers. The proposed model was compared to Gjestvang and Singh (2009) and Singh and Tarray (2014) *RR* models. The proposed procedure was found efficient than the existing *RR* models discussed in their article.

Hussain et al. (2016) suggested two methods to estimate mean and the sensitivity level for the sensitive study variable. The methods originated to be relatively more precise as compared to Gupta et al. (2010) model. The split sample approach and the double sample approach were applied to obtain two *ORR* models. The privacy protection following Zaizai et al. (2009) was also tested to measure the privacy of the respondent's response and the proposed model was more protective as compare to Gupta et al. (2010) model.

To estimate mean and sensitivity level Hussain and Al-Zahrani (2016) suggested the *ORR* model using two different approaches. The two approaches used are the split sample approach and the double sample approach. They followed the assumptions of Gupta et al. (2006) *ORR* model. The model is given as

$$Z_{32(i)} = (1-\alpha)Y + \alpha\left(Y + S_{(i)}\right), \qquad \text{where } i = 1, 2. \qquad (2.4.68)$$

Under the two different approach, the suggested *ORR* model observed to be more efficient as compared to Gupta et al. (2006) model and some other existing models. It was noticed while using double sample approach the privacy issues of the respondent raises but overall, the split sample approach is applicable when a sensitive survey is conducted.

2.5 SOME ESTIMATORS TO ESTIMATE POPULATION MEAN OF SENSITIVE STUDY VARIABLE

Sousa et al. (2010) first introduced ratio estimator to estimate population mean of sensitive study variable using non-sensitive auxiliary variable based on additive model. The ratio estimator proposed under the model $Z_4 = Y + S$, it is due to Pollock and Bek (1971) additive model and used z_4 accordingly. The ordinary mean estimator and its variance using Z_4 are presented respectively as,

$$t_{37} = \bar{z}_4, \qquad (2.5.1)$$

and

$$Var(t_{37}) = \theta\left(S_y^2 + S_S^2\right). \tag{2.5.2}$$

The ratio estimator for sensitive variable *y* consuming Z_4 presented by Sousa et al. (2010) is taken as,

$$t_{38} = \bar{z}_4\left[\frac{\bar{X}}{\bar{x}}\right]. \tag{2.5.2}$$

The bias and the mean square error of t_{38} are as follows,

$$Bias(t_{38}) = \theta\bar{Y}\left(C_x^2 - \rho_{zx}C_xC_z\right), \tag{2.5.3}$$

and

$$MSE(t_{38}) = \theta\bar{Y}^2\left(C_z^2 + C_x^2 - 2\rho_{zx}C_xC_z\right). \tag{2.5.4}$$

The ratio estimator for sensitive variable *y* is said to be efficient under the condition $\rho_{zx} > \frac{1}{2}\frac{C_x}{C_z}$. A transformed ratio estimator was also discussed by Sousa et al. (2010) is presented as,

$$t_{39} = \bar{z}_4\left[\frac{a\bar{X} + b}{a\bar{x} + b}\right], \tag{2.5.5}$$

where *a* and *b* are the unit-free parameters, may be assumed as quantities i.e., coefficient of Kurtosis or coefficient of skewness for non-sensitive variable *X*.

The bias and the mean square errors are as follows,

$$Bias(t_{39}) = \theta\bar{Y}\left(\kappa^2C_x^2 - \kappa\rho_{zx}C_xC_z\right), \tag{2.5.6}$$

and

$$MSE(t_{39}) = \theta\bar{Y}^2\left(C_z^2 + \kappa^2C_x^2 - 2\kappa\rho_{zx}C_xC_z\right), \tag{2.5.7}$$

where $\kappa = \frac{a\bar{X}}{a\bar{X} + b}$. The transformed ratio estimator said to be efficient if the condition $\rho_{zx} > \frac{(\kappa + 1)}{2}\frac{C_x}{C_z}$ is satisfied. The simulation study was performed and determined that both proposed estimators were efficient than the usual unbiased estimator. Also, comparing usual ratio estimator with transformed estimator, it was noticed that the transformed ratio estimator obtain minimal gain than usual ratio estimator.

Gupta et al. (2012) introduced regression estimator and generalized regression-cum-ratio estimator following the underlying assumption of Sousa et al. (2010) that the study variable is sensitive in nature but auxiliary variable is non-sensitive. The expressions of the bias and the *MSE* were derived up to the first order of approximation. The proposed estimators perform better than ratio and usual mean estimator even for modest correlation between sensitive study variable and non-sensitive auxiliary variable. The regression estimator for sensitive study variable is as,

$$t_{40} = \bar{z}_4 + b_{zx}\left(\bar{X} - \bar{x}\right), \tag{2.5.8}$$

where b_{zx} is a constant. Replacing $\bar{z}$ to $\bar{y}$, the above estimator becomes Cochran (1940) regression estimator for non-sensitive study variable *y*. The bias and the *MSE* of the estimator t_{33} are given by,

$$Bias\left(t_{40}\right) = -\theta b_{zx}\left[\frac{\mu_{12}}{\mu_{11}} - \frac{\mu_{03}}{\mu_{02}}\right], \tag{2.5.9}$$

and

$$MSE\left(t_{40}\right) = \theta\bar{Y}^2 C_z^2\left(1-\rho_{zx}^2\right). \tag{2.5.10}$$

Another estimator by Gupta et al. (2012) is written as

$$t_{41} = \left[u_1\bar{z}_4 + u_2\left(\bar{X} - \bar{x}\right)\right]\left[\frac{\bar{X}}{\bar{x}}\right], \tag{2.5.11}$$

where u_1 and u_2 are constants. The estimator t_{40} becomes Ray and Singh (1981) ratio in regression estimator if we replace $\bar{z}$ to $\bar{y}$ and assume that *y* is non-sensitive variable of interest. The expressions of the bias and MSE of the proposed estimator are as,

$$Bias\left(t_{41}\right) = \left(u_1 - 1\right)\bar{Y} + u_1\bar{Y}\theta\left[C_x^2 - \rho_{zx}C_zC_x\right] + u_2\bar{X}\theta C_x^2, \tag{2.5.12}$$

and

$$MSE\left(t_{41}\right) = \frac{\bar{Y}^2 C_z^2\left(1-\rho_{zx}^2\right)\theta\left(1-\theta C_x^2\right)}{C_z^2\left(1-\rho_{zx}^2\right)\theta + \left(1-\theta C_x^2\right)}. \tag{2.5.13}$$

The work of Sousa et al. (2010) and Gupta et al. (2012) was extended by Sousa et al. (2014) from simple random sampling to stratified random sampling design. They assumed that the study variable is sensitive whereas auxiliary variable is non-sensitive in nature. Sousa et al. (2014) presented a combined ratio and a combined regression estimator using Z_4 as,

$$t_{42} = \bar{z}_{st} \frac{\bar{X}}{\bar{x}_{st}}, \tag{2.5.14}$$

and

$$t_{43} = \bar{z}_{st} + b_c\left(\bar{X} - \bar{x}_{st}\right), \tag{2.5.15}$$

where $b_c = \dfrac{\sum_{h=1}^{L} P_h^2 \gamma_h S_{zxh}}{\sum_{h=1}^{L} P_h^2 \gamma_h S_{xh}^2}$ is the sample regression coefficient between Z and X and is scrambled on Y. The bias of the combined ratio and regression estimators are respectively as,

$$Bias\left(t_{42}\right) = \bar{Y} \sum_{h=1}^{L} P_h^2 \gamma_h \left(C_{xh}^2 - C_{zxh}\right), \tag{2.5.16}$$

and

$$Bias\left(t_{43}\right) = \bar{Y} \sum_{h=1}^{L} P_h^2 \gamma_h b_c \left(\frac{\mu_{12h}}{\mu_{11h}} - \frac{\mu_{03h}}{\mu_{02h}}\right). \tag{2.5.17}$$

The expressions of the mean square error of the proposed estimators t_{41} and t_{42} are given respectively as,

$$MSE\left(t_{42}\right) = \bar{Y}^2 \sum_{h=1}^{L} P_h^2 \gamma_h \left(C_{zh}^2 + C_{xh}^2 - 2C_{zxh}\right), \tag{2.5.18}$$

and

$$MSE\left(t_{43}\right) = \bar{Y}^2 \sum_{h=1}^{L} P_h^2 \gamma_h C_{zh}^2 \left(1 - \rho^2{}_{zxh}\right). \tag{2.5.19}$$

They inferred that the suggested estimators t_{41} and t_{42} performs better than usual mean estimator and by using auxiliary information, the accuracy of the estimates improved. Also, Koyuncu et al. (2014) introduced some exponential-type estimators under the assumptions given by Sousa et al. (2010). Koyuncu et al. (2014) used non-sensitive auxiliary information for cases when one or two non-sensitive auxiliary variables are available. The proposed estimator is presented as

$$t_{44}=\left[w_1\bar{z}_4+w_2\left(\bar{X}-\bar{x}\right)\right]\exp\left[\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right]. \tag{2.5.20}$$

The bias and the *MSE* of t_{44} up to the first order of approximation are as,

$$Bias\left(t_{44}\right)=\left(w_1-1\right)\bar{Y}+\theta\left[\frac{1}{2}w_1\bar{Y}\left(\frac{3}{4}C_x^2-C_{zx}\right)+\frac{1}{2}w_2\bar{X}C_x^2\right], \tag{2.5.21}$$

and

$$MSE\left(t_{44}\right)=\bar{Y}^2\left[1-\frac{\theta^2C_x^2\left(\frac{1}{16}C_x^2+C_z^2\left(1-\rho_{zx}^2\right)\right)+4}{4\left(1+C_z^2\theta^2\left(1-\rho_{zx}^2\right)\right)}\right] \tag{2.5.22}$$

The *MSE* of t_{44} is minimum for

$$w_{1(opt)}=\frac{1-\frac{1}{8}\theta C_x^2}{1+\theta C_z^2\left(1-\rho_{zx}^2\right)},$$

and

$$w_{2(opt)}=\frac{\bar{Y}}{2\bar{X}}\frac{C_x^2-2C_x^2+2C_{zx}+\theta C_x^2\left(C_z^2\left(1-\rho_{zx}^2\right)+\frac{1}{4}\left(C_x^2-C_{zx}\right)\right)}{C_x^2\left[1+\theta C_z^2\left(1-\rho_{zx}^2\right)\right]}.$$

The other generalized regression-cum-exponential estimator suggested by Koyuncu et al. (2014) for the case having two non-sensitive auxiliary variables is given by,

$$t_{45}=\left[u_1\bar{z}_4+u_2\left(\bar{X}_1-\bar{x}_1\right)+u_3\left(\bar{X}_2-\bar{x}_2\right)\right]\exp\left[\frac{\left(\bar{X}_1-\bar{x}_1\right)+\left(\bar{X}_2-\bar{x}_2\right)}{\left(\bar{X}_1+\bar{x}_1\right)+\left(\bar{X}_2+\bar{x}_2\right)}\right]. \tag{2.5.23}$$

The bias and the *MSE* of t_{45} are derived as,

$$Bias\left(t_{45}\right)=\left[\left(u_1-1\right)\bar{Z}+\frac{u_1\theta\bar{Z}}{2\left(\bar{X}_1+\bar{X}_2\right)}\left(-\bar{X}_1C_{zx_1}-\bar{X}_2C_{zx2}+\frac{3\bar{X}_1^{\,2}}{4\left(\bar{X}_1+\bar{X}_2\right)}C_{x_1}^2\right.\right.$$

$$+\frac{3\bar{X}_2^{\,2}}{4\left(\bar{X}_1+\bar{X}_2\right)}C_{x_2}^2+\frac{3\bar{X}_1\bar{X}_2}{2\left(\bar{X}_1+\bar{X}_2\right)}C_{x1x2}\Bigg)$$

$$+\frac{u_2\theta\bar{X}_1}{2\left(\bar{X}_1+\bar{X}_2\right)}\left(\bar{X}_1C_{x_1}^2+\bar{X}_2C_{x1x2}\right)$$

$$\frac{u_3\theta\bar{X}_2}{2\left(\bar{X}_1+\bar{X}_2\right)}\left(\bar{X}_2C_{x_2}^2+\bar{X}_1C_{x1x2}\right)\Bigg], \qquad (2.5.24)$$

and

$$MSE\left(t_{45}\right)=\Big[\bar{Z}^2+u_1A-u_2B-u_3C+u_1^2D+u_2^2\bar{X}_1^{\,2}\theta C_{x1}^2$$
$$+u_3^2\bar{X}_2^{\,2}\theta C_{x2}^2+2u_1u_2F+2u_1u_3G+2u_2u_3\bar{X}_1\bar{X}_2\theta C_{x1x2}\Big], \qquad (2.5.25)$$

where

$$A=\bar{Z}^2\left(-2+\theta\left\{\frac{\bar{X}_1C_{zx_1}}{\left(\bar{X}_1+\bar{X}_2\right)}+\frac{\bar{X}_2C_{zx_2}}{\left(\bar{X}_1+\bar{X}_2\right)}-\frac{3\bar{X}_1^2C^2{}_{x_1}}{4\left(\bar{X}_1+\bar{X}_2\right)^2}\right.\right.$$
$$\left.\left.-\frac{6\bar{X}_1\bar{X}_2C_{x_1x_2}}{4\left(\bar{X}_1+\bar{X}_2\right)^4}-\frac{3\bar{X}_2^2C^2{}_{x_2}}{4\left(\bar{X}_1+\bar{X}_2\right)^2}\right\}\right),$$

$$B=\theta\frac{\bar{Z}}{\left(\bar{X}_1+\bar{X}_2\right)}\left(\bar{X}_1^2C^2{}_{x_1}+\bar{X}_1\bar{X}_2C_{x_1x_2}\right),$$

$$C=\theta\frac{\bar{Z}}{\left(\bar{X}_1+\bar{X}_2\right)}\left(\bar{X}_2^2C^2{}_{x_2}+\bar{X}_1\bar{X}_2C_{x_1x_2}\right),$$

$$D=\bar{Z}^2\left(1+\theta\left\{C_z^2+\frac{\bar{X}_1^2C^2{}_{x_1}}{\left(\bar{X}_1+\bar{X}_2\right)^2}+\frac{\bar{X}_2^2C^2{}_{x_2}}{\left(\bar{X}_1+\bar{X}_2\right)^2}-2\frac{\bar{X}_1C_{zx_1}}{\left(\bar{X}_1+\bar{X}_2\right)}\right.\right.$$
$$\left.\left.-2\frac{\bar{X}_2C_{zx_2}}{\left(\bar{X}_1+\bar{X}_2\right)}+2\frac{\bar{X}_1\bar{X}_2C_{x_1x_2}}{\left(\bar{X}_1+\bar{X}_2\right)^4}\right\}\right),$$

$$F=\theta\bar{Z}\left(\frac{\bar{X}_1^2C^2{}_{x_1}}{\left(\bar{X}_1+\bar{X}_2\right)}-\bar{X}_1C_{zx_1}+\frac{\bar{X}_1\bar{X}_2C_{x_1x_2}}{\left(\bar{X}_1+\bar{X}_2\right)}\right),$$

$$G=\theta\bar{Z}\left(\frac{\bar{X}_2^2C^2{}_{x_2}}{\left(\bar{X}_1+\bar{X}_2\right)}-\bar{X}_2C_{zx_2}+\frac{\bar{X}_1\bar{X}_2C_{x_1x_2}}{\left(\bar{X}_1+\bar{X}_2\right)}\right).$$

The minimum $MSE\left(t_{45}\right)$ is given as

$$MSE(t_{45}) = \bar{Z}^2 - \frac{A^2}{4D} - \frac{1}{4D} \frac{\begin{array}{c}(AG-CD)^2\left(D\lambda S_{x1}^2 - F^2\right) \\ +(AF-BD)^2\left(D\lambda S_{x2}^2 - G^2\right) \\ -2(AG-CD)(AF-BD)\left(D\lambda S_{x1x2} - FG\right)\end{array}}{\left(D\lambda S_{x1}^2 - F^2\right)\left(D\lambda S_{x2}^2 - G^2\right) - \left(D\lambda S_{x1x2} - FG\right)^2}. \quad (2.5.26)$$

The simulation study and numerical study were done to examine the efficiency of the proposed estimator by comparing them with Sousal et al. (2010) and Gupta et al. (2012) estimators.

Kalucha et al. (2015) developed additive ratio estimator under the scheme of *SRSWOR* for sensitive surveys. They adapted Gupta et al. (2010) optional randomized response model assumption for the estimation of population mean. The estimator proposed by Kalucha et al. (2015) is given as,

$$t_{46} = \frac{\left(\lambda_2 \bar{z}_{17(1)} - \lambda_1 \bar{z}_{17(2)}\right)}{\left(\lambda_2 - \lambda_1\right)} \left(\frac{\bar{X}_1}{\bar{x}_1} + \frac{\bar{X}_2}{\bar{x}_2}\right)\left(\frac{1}{2}\right). \quad (2.5.27)$$

The derived bias and *MSE* of t_{46} are respectively as,

$$Bias(t_{46}) = \theta_1 \left[\frac{\bar{Y}}{2} C_x^2 - \frac{\lambda_2}{\lambda_2 - \lambda_1} \frac{\rho_{xy}\sigma_y C_x}{2}\right] + \theta_2 \left[\frac{\bar{Y}}{2} C_x^2 - \frac{\lambda_2}{\lambda_2 - \lambda_1} \frac{\rho_{xy}\sigma_y C_x}{2}\right], \quad (2.5.28)$$

and

$$MSE(t_{46}) = \theta_1 \left[\left(\frac{\lambda_2}{\lambda_2 - \lambda_1}\right)^2 \sigma_{z_1}^2 + \frac{\bar{Y}^2}{4} C_x^2 - \bar{Y}\left(\frac{\lambda_2}{\lambda_2 - \lambda_1}\right)\rho_{xy}\sigma_y C_x\right]$$
$$+\theta_2 \left[\left(\frac{\lambda_2}{\lambda_2 - \lambda_1}\right)^2 \sigma_{z_1}^2 + \frac{\bar{Y}^2}{4} C_x^2 + \bar{Y}\left(\frac{\lambda_2}{\lambda_2 - \lambda_1}\right)\rho_{xy}\sigma_y C_x\right]. \quad (2.5.29)$$

The proposed estimator was than compared mathematically and numerically with the ordinary *RR* mean estimator t_{37}.

Gupta et al. (2016) proposed some improved exponential type estimators under simple random sampling without replacement (*SRSWOR*) based on *SRR*. The suggested estimator is as,

$$t_{47} = [u_3 \bar{z}_4 + u_4] \exp\left[\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right], \qquad (2.5.30)$$

where u_3 and u_4 are constants. The bias and the *MSE*'s are respectively given as,

$$Bias(t_{47}) = (u_3 - 1)\bar{Y} + u_3 \bar{Y} \theta \left[\frac{3}{8} C_x^2 - \frac{1}{2} \rho_{zx} C_z C_x\right] + u_4 \left[1 + \frac{3}{8} \theta C_x^2\right], \qquad (2.5.31)$$

and

$$MSE(t_{47}) = \bar{Y}^2 \left[1 - \frac{\begin{array}{c}\bar{Y}^2 \theta C_z^2 + \frac{3}{4} \theta^2 C_z^2 C_x^2 \left(1 - \rho_{zx}^2\right) \\ \left(1 + \frac{3}{16} \theta C_x^2\right) + \frac{1}{64} \theta C_x^2 \left(\theta C_z C_x \rho_{zx}\right)^2\end{array}}{\theta C_z^2 + \theta^2 C_z^2 C_x^2 \left(1 - \rho_{zx}^2\right)}\right]. \qquad (2.5.32)$$

Another generalized exponential type estimators using two non-sensitive auxiliary variables is written as,

$$t_{48} = [u_5 \bar{z}_4 + u_6] \exp\left[\frac{\bar{X}_1 - \bar{x}_1}{\bar{X}_1 + \bar{x}_1} + \frac{\bar{X}_2 - \bar{x}_2}{\bar{X}_2 + \bar{x}_2}\right]. \qquad (2.5.33)$$

The mean square error of the above estimator is given by,

$$MSE(t_{48}) = \bar{Y}^2 \left[1 - \frac{B^* C^* + A^* D^{*2} - 2 C^* D^* E^*}{A^* B^* - E^{*2}}\right]. \qquad (2.5.34)$$

where

$$A^* = 1 + \theta C_z^2 + \theta C_{x_1}^2 + \theta C_{x2}^2 + \theta C_{x_1} C_{x2} \rho_{x_1 x_2} - 2\theta C_{x_1} C_z \rho_{zx_1} - 2\theta C_{x2} C_z \rho_{zx_2},$$

$$B^* = 1 + \theta C_{x_1}^2 + \theta C_{x2}^2 + \theta C_{x_1} C_{x2} \rho_{x_1 x_2},$$

$$C^* = 1 + \frac{3}{8} \theta C_{x_1}^2 + \frac{3}{8} \theta C_{x2}^2 + \frac{1}{4} \theta C_{x_1} C_{x2} \rho_{x_1 x_2} - \frac{1}{2} \theta C_{x_1} C_z \rho_{zx_1} - \frac{1}{2} \theta C_{x2} C_z \rho_{zx_2},$$

$$D^* = 1 + \frac{3}{8} \theta C_{x_1}^2 + \frac{3}{8} \theta C_{x2}^2 + \frac{1}{4} \theta C_{x_1} C_{x2} \rho_{x_1 x_2},$$

$$E^{*}=1+\theta C_{x_1}^{2}+\theta C_{x2}^{2}+\theta C_{x_1}C_{x2}\rho_{x_1x_2}-\theta C_{x_1}C_z\rho_{zx_1}-\theta C_{x2}C_z\rho_{zx_2}.$$

The simulation and numerical study showed that the proposed estimator works efficiently than Koyuncu et al. (2014), usual unbiased mean estimator, ratio estimator, regression estimator and Gupta et al. (2012) estimator.

Mushtaq et al. (2016) first considered the problem of estimating population mean of a sensitive variable under two-phase sampling scheme based on *RR* technique. They presented ratio, regression and general class of estimators for stratified random sampling design. The generalized estimator proposed based on Z_4 model is as,

$$t_{49}=\left[u_7\bar{z}_{st}+u_8\left(\bar{x}'_{st}-\bar{x}_{st}\right)\right]\left[\alpha\left(\frac{\bar{x}'_{st}-b_{st}}{\bar{x}_{st}+b_{st}}\right)+\left(1-\alpha\right)\exp\left(\frac{\bar{x}'_{st}-\bar{x}_{st}}{\left(\bar{x}'_{st}-\bar{x}_{st}\right)+2b_{st}}\right)\right], \tag{2.5.35}$$

where u_7 and u_8 are unknown weights, $\alpha\epsilon\left[0,1\right]$ and b_{st} is the real number. Also, $\bar{x}_{st}{}'$ is the first-phase auxiliary variable under stratified sampling design and the auxiliary variable and reported response at second-phase are respectively $\bar{x}_{st}$ and $\bar{z}_{st}$. The bias and the mean square error of the above estimator are written as,

$$Bias\left(t_{49}\right)=\left(u_7-1\right)\bar{Y}+u_7\bar{Y}\left[\frac{1}{8}g'^{2}\left(3+5\alpha\right)\Lambda_h\vartheta_{02}-\frac{1}{2}g'\left(1+\alpha\right)\Lambda_h\vartheta_{11}\right]$$
$$+u_8\bar{X}\left[\frac{1}{2}g'\left(1+\alpha\right)\Lambda_h\vartheta_{02}\right], \tag{2.5.36}$$

and

$$MSE\left(t_{49}\right)=\bar{Y}^2\Big[\left(u_7-1\right)^2+u_7^2\{\theta_h\vartheta_{20}$$
$$+\Lambda_h\left(\frac{1}{4}g'^2\vartheta_{02}\left(\alpha^2+7\alpha+4\right)-2g'\vartheta_{11}\left(1+\alpha\right)\right)$$
$$-2u_7\Lambda_h\left\{\frac{1}{8}g'^2\vartheta_{02}\left(5\alpha+3\right)-\frac{1}{2}g'\vartheta_{11}\left(1+\alpha\right)\right\}+u_8^2\frac{\bar{X}^2}{\bar{Y}^2}\Lambda_h\vartheta_{02}$$
$$-u_8\frac{\bar{X}}{\bar{Y}}g'\Lambda_h\vartheta_{02}\left(1+\alpha\right)-2u_7u_8\frac{\bar{X}}{\bar{Y}}\Lambda_h\left(\vartheta_{11}-g'\left(1+\alpha\right)\vartheta_{02}\right)\Big]. \quad (2.4.37)$$

where

$$\vartheta_{p,q}=\sum_{h=1}^{L}W_h^{p+q}\frac{E\left(\bar{z}_h-\bar{Z}_h\right)^p E\left(\bar{x}_h-\bar{X}_h\right)^q}{\bar{Z}^p\bar{X}^q}, g'=\frac{\bar{X}}{\bar{X}+b_{st}}\Lambda_h=\theta_h-\theta_h{}',$$

$$\theta_h=\left(\frac{1}{n_h}-\frac{1}{N_h}\right),$$

and

$$\theta'_h=\left(\frac{1}{n'_h}-\frac{1}{N_h}\right).$$

The *MSE* of t_{49} is minimum for the optimum values of u_7 and u_8 as,

$$u_{7(opt)}=\frac{1-\frac{1}{8}\Lambda_h g'^2\left(4\alpha^2+3\alpha+1\right)\vartheta_{02}}{1+\left\{\theta_h V_{20}\left(1-\rho_{zxh}^2\right)-g'^2\frac{1}{4}\left(\alpha+3\alpha^2\right)\Lambda_h\vartheta_{02}\right\}},$$

and

$$u_{8(opt)}=\frac{\bar{Y}}{\bar{X}}\left\{\frac{1}{2}g\left(1+\alpha\right)+k_1\left(\frac{V_{11}}{V_{02}}-g\left(1+\alpha\right)\right)\right\}.$$

The simulation study was performed to check the efficiency of the proposed model with some available estimators.

Motivated by Kalucha et al. (2015) ratio estimator, Gupta et al. (2017) presented a non-optional regression type estimator to estimate population mean. The proposed estimator was made for the assumption using two-non-sensitive auxiliary variable.

$$t_{50}=\frac{\left(\lambda_{(2)}\bar{z}_{17(1)}-\lambda_{(1)}\bar{z}_{17(2)}\right)}{\left(\lambda_{(2)}-\lambda_{(1)}\right)}\left\{b_{z_1x_1}\left(\bar{X}_1-\bar{x}_1\right)+b_{z_2x_2}\left(\bar{X}_2-\bar{x}_2\right)\right\}\left(\frac{1}{2}\right), \quad (2.5.38)$$

where $b_{z_1x_1}$ and $b_{z_2x_2}$ are the sample regression coefficients between z_i and x_i respectively. The bias and the mean square of the estimator t_{50} is written by,

$$Bias(t_{50}) = \left\{-\frac{1}{2}\theta_1 b_{z_1x} - \frac{1}{2}\theta_2 b_{z_2x}\right\}\left[\frac{\mu_{12}}{\mu_{11}} - \frac{\mu_{03}}{\mu_{02}}\right], \tag{2.5.39}$$

and

$$MSE(t_{50}) = \frac{1}{(\lambda_2 - \lambda_1)^2}\left[\lambda_2^2\theta_1\sigma_{z_1}^2 + \lambda_1^2\theta_2\sigma_{z_2}^2\right] + \frac{\rho_{YX}^2\sigma_Y^2}{4}\alpha - \rho_{YX}^2\sigma_Y^2\beta \tag{2.5.40}$$

It was verified that the proposed estimator is efficient as compared to ordinary *ORR* mean estimator and Kalucha et al. (2015) ratio type estimator. The sensitivity level (*W*) was calculated through simulation using split sample approach.

2.6 SOME EXISTING ESTIMATORS TO ESTIMATE POPULATION VARIANCE

The consideration above is provided for the estimation of population mean for sensitive and non-sensitive variable of interest. But there is a situation comes when one needs to measure the variation between two different goods or climate factors of different places, etc. for such a situation we estimate population variance of study variable. The conventional unbiased variance estimator for the estimation of non-sensitive study variable '*y*' is as,

$$t_{51} = s_y^2. \tag{2.6.1}$$

The variance of the unbiased estimator t_{51} is given by

$$Var(t_{51}) = \theta S_y^4 \delta_{40}, \tag{2.6.2}$$

where $S_y^4 = \frac{1}{(N-1)}\sum_{i=1}^{N}(Y_i - \bar{Y})^4$, and $\delta_{uv} = \frac{1}{(N-1)}\sum_{i=1}^{N}(Y_i - \bar{Y})^u (X_i - \bar{X})^v$.

The ratio estimator to estimate population variance of non-sensitive study variable was introduced by Isaki (1983) for the situation when auxiliary information (*x*) is available and the correlation between *x* and *y* is positive. His proposed ratio estimator is as

$$t_{52} = s_y^2\left[\frac{S_x^2}{s_x^2}\right]. \qquad (2.6.3)$$

The bias and the *MSE* of the estimator t_{52} are given respectively,

$$Bias(t_{52}) = \theta S_y^2(\delta_{04} - \delta_{22}), \qquad (2.6.3)$$

and

$$MSE(t_{52}) = \theta S_y^4[\delta_{40} + \delta_{04} - 2\delta_{22}]. \qquad (2.6.4)$$

Also, a regression estimator for the estimation of variance was suggested by Isaki (1983) when *x* and *y* are linearly related to each other, the unbiased estimator is written as,

$$t_{53} = s_y^2 + b_{yx}\left(S_x^2 - s_x^2\right), \qquad (2.6.5)$$

where $b_{yx} = \frac{s_y^2\delta_{22}^2}{s_x^2\delta_{04}}$ is the sample regression coefficient. The *MSE* of t_{53} is presented as,

$$MSE(t_{53}) = \theta S_y^4\left(\delta_{40} - \frac{\delta_{22}^2}{\delta_{04}}\right). \qquad (2.6.6)$$

Kadilar and Cingi (2006) introduced a ratio-type estimator to estimate population variance for the non-sensitive study variable *y* as,

$$t_{54} = w_1 s_y^2 + w_2\left[s_y^2\frac{S_x^2}{s_x^2}\right]\tau, \qquad (2.6.7)$$

where $\tau = \frac{1+\theta\delta_{22}}{1+\theta\delta_{04}}$ and $w_1 + w_2 = 1$. The bias and the *MSE* of the t_{54} are respectively as,

$$Bias(t_{54}) = S_y^2\left[(w_1 - 1) + w_2\tau\theta\frac{\delta_{04}}{\delta_{02}^2}\right]. \qquad (2.6.8)$$

and

$$MSE\left(t_{54}\right)=\theta S_y^4\left[d^2\left(\frac{\delta_{40}}{\delta_{20}^2}\right)-2w_2 d\tau\delta_{22}+w_2^2\tau^2\left(\frac{\delta_{04}}{\delta_{02}^2}\right)\right], \qquad (2.6.9)$$

where $d=w_1+w_2\tau$. The *MSE* of t_{54} is minimum for

$$w_1=\frac{(\tau-1)\left(\frac{\delta_{40}}{\delta_{20}^2}\right)+\delta_{22}(1-2\tau)+\left(\frac{\delta_{04}}{\delta_{02}^2}\right)\tau}{\left[(1-\tau)^2/\tau\right]\left(\frac{\delta_{40}}{\delta_{20}^2}\right)+2\delta_{22}(1-\tau)+\left(\frac{\delta_{04}}{\delta_{02}^2}\right)\tau}, \text{ and } w_2=1-w_1.$$

The estimator suggested by Kadilar and Cingi (2006) performed better than Isaki (1983) regression and ratio estimators under certain conditions and when $\tau=1$, the above estimator is equally efficient as Isaki's (1983) regression estimator.

Singh et al. (2011) presented an exponential product estimator to estimate population variance for non-sensitive study variable and the estimator proposed is given by,

$$t_{55}=s_y^2\exp\left[\frac{s_x^2-S_x^2}{s_x^2+S_x^2}\right]. \qquad (2.6.10)$$

The bias and the *MSE* of the estimator t_{55} are given by,

$$Bias\left(t_{55}\right)=\theta S_y^2\left(\frac{\delta_{04}}{8}+\frac{\delta_{22}}{2}-\frac{5}{8}\right), \qquad (2.6.11)$$

and

$$MSE\left(t_{55}\right)=\theta S_y^4\left[\delta_{40}+\frac{\delta_{04}}{4}+\delta_{22}-\frac{9}{4}\right]. \qquad (2.6.12)$$

A generalized exponential estimator suggested by Yadav and Kadilar (2013) is written as,

$$t_{56}=s_y^2\exp\left[\frac{S_x^2-s_x^2}{S_x^2+(\alpha-1)s_x^2}\right]. \qquad (2.6.13)$$

The bias and the *MSE* of the estimator t_{56} are given by,

$$Bias(t_{56}) = \theta S_y^2 \left(\frac{\delta_{04}}{2\alpha^2} \left(2\alpha(1-\kappa) - 1 \right) \right), \tag{2.6.14}$$

and

$$MSE(t_{56}) = \theta S_y^4 \left[\delta_{40} + \frac{\delta_{04}}{\alpha^2} + (1 - 2\alpha\kappa) \right]. \tag{2.6.15}$$

where $\kappa = \frac{\delta_{22}}{\delta_{04}}$ and the $MSE(t_{56})$ is minimum for $\alpha = \frac{1}{\kappa}$. The estimator proposed by Yadav and Kadilar (2013) was found to be more efficient than usual variance estimator and ratio estimator.

Asghar et al. (2014) introduced generalized estimators to estimate population variance of study variable Y using mean auxiliary variables is given as,

$$t_{57} = \tau' s_y^2 \exp\left[\alpha^* \left(\frac{\bar{X} - \bar{x}}{\bar{X} + (a^* - 1)\bar{x}} \right) \right]. \tag{2.6.16}$$

The bias and the *MSE* of t_{57} are written as,

$$Bias(t_{57}) = \theta S_y^2 \left[\tau' \left\{ 1 + \frac{1}{2} \nu'^2 C_x^2 - \nu' \delta_{21} C_x \right\} \right] - S_y^2, \tag{2.6.17}$$

and

$$MSE(t_{57}) = \theta S_y^4 \left[\tau'^2 \left\{ 1 + \delta_{40} - 2\alpha' \delta_{21} C_x + \alpha'^2 \right\} + (1 - 2\tau') \right]. \tag{2.6.18}$$

where $\nu' = \frac{\alpha^*}{a^*}$. The $MSE(t_{57})$ is minimum for the optimal values of ν' and τ' are presented respectively as $\nu' = \delta_{21}(C_x)^{-1}$ and $\tau' = \left(\delta_{40} - \delta_{21}^2\right)^{-1}$. The estimator t_{57} works efficiently as compare to usual unbiased estimator t_{51}, Isaki (1983) estimator and Yadav and Kadilar (2013) generalized estimators.

Sanaullah et al. (2016) developed generalized exponential estimator for the estimation of population variance using two auxiliary variables x_1 and x_2. They presented estimator for two-phase sampling considering unknown population means of auxiliary variable is as follows,

$$t_{58} = \eta s_y^2 \exp\left[\alpha_1\left(\frac{a_1\bar{x}_1'}{(a_1-1)\bar{x}_1' + \bar{x}_1} - 1\right)\right]\exp\left[\alpha_2\left(\frac{a_2\bar{x}_2'}{(a_2-1)\bar{x}_2' + \bar{x}_2} - 1\right)\right]. \tag{2.6.19}$$

The bias and the *MSE* of t_{58} are written as,

$$Bias(t_{58}) = S_y^2\left[\eta\left\{\begin{array}{l}1+\left(\frac{\alpha_1}{a_1^2}+\frac{\alpha_1^2}{2a_1^2}\right)\theta_2 C_{x1}^2+\left(\frac{\alpha_2}{a_2^2}+\frac{\alpha_2^2}{2a_2^2}\right)\theta_2 C_{x2}^2 \\ -\theta_2\frac{\alpha_1}{a_1}\delta_{210}C_{x1}-\theta_2\frac{\alpha_2}{a_2}\delta_{201}C_{x2}+\theta_2\frac{\alpha_1\alpha_2}{a_1 a_2}C_{x1}C_{x2}\rho_{x1x2}\end{array}\right\}-1\right], \tag{2.6.20}$$

and

$$MSE(t_{58}) = S_y^4\left[1+\eta^2\left\{\begin{array}{l}1+\frac{\delta_{400}}{n_2}+\theta_2\nu_1^2C_{x1}^2+\theta_2\nu_2^2C_{x2}^2-2\theta_2\nu_1\delta_{210}C_{x1} \\ -2\theta_2\nu_2\delta_{201}C_{x2}+2\theta_2\nu_1\nu_2\rho_{x1x2}C_{x1}C_{x2}\end{array}\right\}-2\eta\right]. \tag{2.6.21}$$

The $MSE(t_{58})$ is minimized for the optimum values of ν_1, ν_2 and η are as,

$$\nu_1 = (\delta_{210} - \delta_{201}\rho_{x1x2})\left(C_{x1}\left(1-\rho_{x1x2}^2\right)\right)^{-1},$$

$$\nu_1 = (\delta_{201} - \delta_{210}\rho_{x1x2})\left(C_{x2}\left(1-\rho_{x1x2}^2\right)\right)^{-1}$$

and

$$\nu_1 = \left(1+\frac{\delta_{400}}{n_2}+\theta_2\left(\begin{array}{l}\nu_1^2C_{x1}^2+\nu_2^2C_{x2}^2-2\nu_1\delta_{210}C_{x1} \\ \qquad -2\nu_2\delta_{201}C_{x2}+2\nu_1\nu_2\rho_{x1x2}C_{x1}C_{x2}\end{array}\right)\right).$$

They used two populations to examine the performance of the generalized exponential estimator and found to be better estimate than some existing estimators discussed in their paper.

Many survey researchers paid attention to contribute for the estimation of population variance for non-sensitive study variable such as Singh and Salonki (2013), Khan and Shabbir (2013), Yadav and Misra (2015) and Asghar et al. (2017).

Singh et al. (2015) first noticed the problem of estimation population variance of sensitive study variable. They used scrambled response mechanism and presented regression estimator. The *SRR* model used for the study was Eichhorn and Hayre (1983) multiplicative model. The estimator presented is as,

$$t_{58} = s_y^{*2} + \hat{B}\left(S_x^2 - s_x^2\right), \tag{2.6.22}$$

where $s_y^{*2} = \dfrac{1}{\theta^2}\left[s_z^2 - \dfrac{S_s^2}{n\left(S_s^2 - \bar{S}^2\right)}\sum_{i=1}^{n} Z_i^2\right]$ and $\hat{B} = \dfrac{\delta_{22} - s_y^{*2}s_x^2}{\delta_{04} - s_x^4}$.

The variance of the estimator t_{58} are given by,

$$Var\left(t_{58}\right) = Var\left(s_y^{*2}\right) + \frac{\left(S_x^2 - s_x^2\right)^2}{\left(\delta_{04} - s_x^4\right)^2} Var\left(\delta_{22} - s_y^{*2}s_x^2\right)$$

$$+2\frac{\left(S_x^2 - s_x^2\right)^2}{\left(\delta_{04} - s_x^4\right)^2} Cov\left(s_y^{*2}, \delta_{22} - s_y^{*2}s_x^2\right). \tag{2.6.23}$$

CHAPTER 3

PROPOSED SCRAMBLED RANDOMIZED RESPONSE MODELS

3.1 INTRODUCTION

In this chapter, some generalized *SRR* models have been proposed using scrambling device. The idea is to propose more flexible *SRR* models in terms of accuracy and privacy protection as compared to some *SRR* models present in literature. The *SRR* models proposed have been presented for the situation when two different scrambling variables are used. The coding mechanism of the response on Y to estimate $\bar{Y}$ and the *SRR* models proposed are convex combination of additive and multiplicative models.

3.2 PROPOSED PROCEDURE

The procedure for the proposed models is as follows:

Let a finite population of N units and Y_i be a sensitive quantitative variable of interest with unknown population mean and variance respectively as $\bar{Y}$ and S_y^2. Let S_i and R_i be the values of random variables with known distribution and are independent of the sensitive variable Y_i. Usually the respondent does not reveal the true response Y_i, the random variables S_i and R_i are used to scramble the sensitive study variable y_i.

The expectation and variance with respect to design based are denoted as E_D and V_D, and the model-based expectation and variance using *SRR* device are respectively as E_M and V_M. But the overall expectation and variance are considered as $E = E_M E_D$ and $V = V_M V_D$. Therefore, the following assumptions are made as,

$$\begin{aligned} &E_D\left(Y\right)=\bar{Y}, E_M\left(S\right)=\bar{S}=0, E_M\left(R\right)=\bar{R}=1, \\ &V_D\left(Y\right)=S_Y^2, V_M\left(S\right)=S_S^2, V_M\left(R\right)=S_R^2. \end{aligned} \qquad (3.2.1)$$

3.3 PROPOSED SRR MODEL-I

Motivated from Diana and Perri (2010), a new *SRR* model is introduced for extreme sensitive surveys using scrambling device to provide more protection to the respondents. The reported response is given as,

$$Z_{G1} = gYR + (1-g)(Y + aS), \tag{3.3.1}$$

where $g \in [0,1)$ and $a \in [-1,1)$ are suitable constants controlled by the researcher. The above reported response is the combination of three *SRR* models considering different choices of *g* and *a* are as follows:

i) If $g = 0$ and $a = 1$, the Z_{G1} becomes Pollock and Bek's (1976) additive

ii) If $g = 0$ and $a = -1$, the model becomes the subtractive *SRR* model.

iii) If $g = 0$ and $a = 0$, gives true response *Y*.

iv) If $g = 1$, the model reduces to Eichhorn and Hayre (1983) multiplicative *SRR* model.

The resultant mean, variance respectively for the *SRR* model Z_{G1} are recognized as,

$$E(Z_{G1}) = \bar{Y}, \tag{3.3.2}$$

and

$$Var(Z_{G1}) = S^2_{Z_{G1}} = g^2\left(\theta S_Y^2 + S_R^2 \frac{\sum_{i=1}^{N} Y_i^2}{Nn}\right) + (1-g)^2\left(a^2 \frac{S_S^2}{n} + \theta S_Y^2\right), \tag{3.3.3}$$

where $\theta = \left(\frac{1}{n} - \frac{1}{N}\right)$.

The terms of covariance $\left(\sigma_{XZ_{G1}}\right)$, coefficient of variation (C_Z) and correlation $\rho_{XZ_{G1}}$ are respectively as,

$$\sigma_{XZ_{G1}} = \sigma_{XY}, C^2_{Z_{G1}} = \frac{S^2_{Z_{G1}}}{\bar{Z}^2_{G1}}, \tag{3.3.4}$$

and

$$\rho_{XZ_{G1}} = \rho_{YX} \frac{S_Y}{S_{Z_{G1}}}. \tag{3.3.5}$$

3.4 PROPOSED SRR MODEL-II

Another generalized *SRR* model is introduced for highly sensitive surveys. The reported response using two independent scrambling variables is given as,

$$Z_{G2} = g(YR + aS) + (1-g)YR. \tag{3.4.1}$$

We note that for $g = 0$, the model reduces to Eichhorn and Hayre (1983) multiplicative model and for $g = a = 1$, generates Huang (2010) model. It is relatively clear that the proposed method is the combination of additive and multiplicative approaches in order to gain greater confidence of the respondents regarding their privacy protection.

For the proposed *SRR* model II, the following results are given.

The mean and variance of the model are given as,

$$E_M E_D\left(Z_{G_2}\right) = \bar{Y}, \tag{3.4.2}$$

and

$$Var\left(Z_{G_2}\right) = S^2_{Z_{G2}} = g^2 a^2 \frac{S_S^2}{n} + \theta S_Y^2 + S_R^2 \frac{\sum_{i=1}^{N} Y_i^2}{Nn}. \tag{3.4.3}$$

So the terms of covariance, coefficient of variation and correlation may be written as,

$$\sigma_{XZ_{G2}} = \sigma_{XY}, C^2_{Z_{G2}} = \frac{S^2_{Z_{G2}}}{\bar{Z}^2_{G2}}, \text{ and } \rho_{XZ_G} = \rho_{YX} \frac{S_Y}{S_{Z_{G2}}}. \tag{3.4.4}$$

3.5 PROPOSED SRR MODEL-III

The forth *SRR* proposed model is introduced by following Saha's (2007) approach combining additive and multiplicative *SRR* models for two scrambling variables. The observed response is presented as,

$$Z_{G_3} = g\left(Y + aS\right) + \left(1 - g\right)R\left(Y + aS\right). \tag{3.5.1}$$

The following possibilities can be made for different choices of g's,

i) If $g = 0$ and $a = 1$, the model becomes $Z_{G_3^1} = Y\left(R + S\right)$, which is the combination of multiplicative and additive model presented by Saha (2007).

ii) If $g = 1$ and $a = 1$, Pollock and Bek's (1976) additive model is obtained.

iii) If $g = 0$ and $a = -1$, this gives Hussain (2012) subtractive model.

The mean and the variance of Z_{G3} are given as

$$E_M E_D\left(Z_{G3}\right) = \bar{Y}, \tag{3.5.2}$$

and

$$\begin{aligned} Var\left(Z_{G3}\right) = S^2_{Z_{G3}} &= \theta S_y^2 + a^2 g^2 S_S^2\left(\frac{1}{n} + \frac{1}{N^2}\right) \\ &+ \left(1-g\right)^2\left[\frac{1}{N^2}\sum_{i=1}^{N} Y_i^2 + a^2 S_S^2\right] S_R^2 \\ &+ 2g\left(1-g\right)\frac{1}{N}\left(a^2 S_S^2\left(1 + \frac{1}{n} - \frac{1}{N}\right) + S_y^2\right). \end{aligned} \tag{3.5.3}$$

The covariance, coefficient of variation and correlation terms are presented respectively as

$$\sigma_{XZ_{G3}} = \sigma_{YX}, \; C_{Z_3} = \frac{S_{Z_{G3}}}{\bar{Z}_{G3}}, \text{ and } \rho_{XZ_{G3}} = \rho_{YX}\frac{S_Y}{S_{Z_{G3}}}. \tag{3.5.4}$$

3.6 PROPOSED SRR MODEL-IV

Following Diana and Perri (2010) another generalized *SRR* model for two scrambled variables is proposed. The reported response is as:

$$Z_{G4} = R\left[g\left(Y + aS\right) + \left(1 - g\right)Y\right]. \tag{3.6.1}$$

The following options can be made for different choices of g's:

i) If $g = 1$, the model reduces to Saha (2007) model.

ii) If $g = 0$ or ($g = 1$ and $a = 0$), these choices generates Eichhorn and Hayre (1983) multiplicative model.

The mean and the variance of Z_{G4} are given as

$$E_M E_D\left(Z_{G4}\right) = \bar{Y}. \tag{3.6.2}$$

and

$$\begin{aligned} Var\left(Z_{G4}\right) = S^2_{Z_{G4}} = g^2 &\left[\begin{array}{l}\theta\left(S_y^2 + S_S^2\left(1 + S_R^2\right)\right) \\ \quad + S_R^2\left(\frac{1}{N}a^2 S_S^2 + \frac{1}{n}\sum_{i=1}^{N} Y_i^2 + \frac{1}{N^2}\sum_{i=1}^{N} Y_i^2\right)\end{array}\right] \\ &+ \left(1 - g\right)^2\left[\theta\left(S_y^2 + \frac{1}{N}S_R^2\sum_{i=1}^{N} Y_i^2\right) + \frac{1}{N^2}S_R^2\sum_{i=1}^{N} Y_i^2\right] \\ &+ 4g\left(1 - g\right)\theta\left(S_y^2 + \frac{1}{N}S_R^2\sum_{i=1}^{N} Y_i^2\right). \end{aligned} \tag{3.6.3}$$

The terms of covariance, coefficient of variation and correlation are given as

$$\sigma_{XZ_{G4}} = \sigma_{YX},\ C_{Z4} = \frac{S_{Z_{G4}}}{\bar{Z}_{G4}} \text{ and } \rho_{ZX} = \rho_{YX}\frac{S_Y}{S_{Z_{G4}}}. \tag{3.6.4}$$

3.7 MEASURE OF PRIVACY PROTECTION

The privacy protection measure is used to evaluate the closeness of the true value of the sensitive variable and the respondent's response. Assume a randomized response device in which the respondent is selected in the sample to provide randomized response z_i and a definite value of sensitive variable y_i is also taken. The magnitude of deviation of z_i and y_i is said to be the privacy protection measure of the randomized response device used. The mathematical form presented by Zaizai et al. (2009) is defined as,

$$\Delta = E_M E_D \left(z_i - y_i\right)^2. \tag{3.7.1}$$

3.7.1 Protection Measures of Different SRR Model

This section provides some privacy protection measures adapting (3.7.1) for some of the *SRR* models presented in literature as well as for the proposed *SRR* models.

i) The additive model presented by Pollock and Bek's (1976) is as:

$$Z_4 = Y + S. \tag{3.7.1.1}$$

The privacy protection degree to the respondents following Z_4 is presented as,

$$\Delta_1 = E\left(y_i + s_i - y_i\right)^2 = S_S^2. \tag{3.7.1.2}$$

ii) Eichhorn and Hayre's (1983) multiplicative model given in (2.3.17) is taken as,

$$Z_7 = YR. \tag{3.7.1.3}$$

The privacy protection degree to the respondents using Z_7 is,

$$\Delta_2 = E\left(y_i r_i - y_i\right)^2 = \left(S_Y^2 + \bar{Y}^2\right)S_R^2. \tag{3.7.1.4}$$

iii) The *SRR* model suggested by Diana and Perri (2010) is as follows,

$$Z_{16} = g\left(Y + S\right) + \left(1 - g\right)YR. \tag{3.7.1.5}$$

The measure of privacy protection for the observed response in (3.7.1.5) is specified as,

$$\Delta_3 = E\left(z_{16_i} - y_i\right) = \left(1 - g\right)^2\left(S_Y^2 + \bar{Y}^2\right)S_R^2 + g^2 S_S^2. \tag{3.7.1.6}$$

iv) For the proposed *SRR* model I defined in (3.3.1) is modified for the privacy protection degree as,

$$\Delta_4 = E\left(Z_{G1} - Y\right)^2 = g^2 S_R^2 \left(\bar{Y}^2 + S_Y^2\right) + (1-g)^2 a^2 S_S^2. \qquad (3.7.1.7)$$

v) The proposed *SRR* model II given in (3.4.1) may provide the privacy protection measure as follows,

$$\Delta_5 = E\left(Z_{G2} - Y\right)^2 = S_R^2 \left(\bar{Y}^2 + S_Y^2\right) + g^2 a^2 S_S^2. \qquad (3.7.1.8)$$

vi) The measure of privacy protection using model III is provided as follows,

$$\Delta_6 = E\left(Z_{G3} - Y\right)^2 = (1-g)^2 \left(S_R^2 + S_S^2\right)\left(\bar{Y}^2 + S_Y^2\right) + \left[g^2 + 2g(1-g)\bar{Y}\right] S_S^2. \qquad (3.7.1.9)$$

vii) The protection degree derived using the proposed model IV is as,

$$\Delta_7 = E\left(Z_{G4} - Y\right)^2 = S_R^2 \left(\bar{Y}^2 + S_Y^2\right) + g^2 S_S^2 \left(1 + S_R^2\right). \qquad (3.7.1.10)$$

3.7.2 Comparison of Privacy Protection Measures of Different SRR Models

To examine how much protection does a scrambled model provides, the comparisons among proposed *SRR* models and some existing models discussed in section 3.8.1 are performed. The greater privacy is protected if the value of protection degree higher and this also ascertains that one can obtain more reliable information from the respondents.

The comparisons of the protection degree of different *SRR* models are as follows:

i. $\Delta_4 > \Delta_1$

$$\text{If } \left[g S_R^2 - \frac{a^2 S_S^2 (2-g)}{\left(\bar{Y}^2 + S_y^2\right)} \right] > 0. \qquad (3.7.2.1)$$

The above expression represents that the degree of privacy protection of proposed *SRR* model is higher than that of Additive model.

ii. $\Delta_4 > \Delta_2$

The above condition holds if

$$\left[\frac{a^2 S_S^2 (1-g)^2}{\left(\bar{Y}^2 + S_y^2\right)} - \left(1-g^2\right) S_R^2\right] > 0. \qquad (3.7.2.2)$$

This presents that proposed generalized model I gains higher protection of the respondents as compared to the multiplicative model.

iii. $\Delta_4 > \Delta_3$

$$\left[(1+2g) - \left(g^2 + a^2 (1-g)^2\right) \frac{a^2 S_S^2}{S_R^2 \left(\bar{Y}^2 + S_y^2\right)}\right] > 0. \qquad (3.7.2.3)$$

If the conditions (3.7.2.3) holds, this shows that the privacy protection degree of the proposed model I is higher than Diana and Perri (2010) generalized *SRR* model. Similarly for the proposed *SRR* model II, if the following conditions met, greater privacy will be achieved using proposed model II.

iv. $\Delta_5 > \Delta_1$

If the following condition is met,

$$\left[1 - \frac{\left(1 - a^2 g^2\right) S_S^2}{S_R^2 \left(\bar{Y}^2 + S_y^2\right)}\right] > 0. \qquad (3.7.2.4)$$

The expression in (3.7.2.4) shows that the privacy degree of proposed model III than Himmelfarb and Edgell (1983) additive model.

v. $\Delta_5 > \Delta_2$

If the following condition is met,

$$a^2 g^2 S_S^2 > 0. \tag{3.7.2.5}$$

It is always true for $0 < g < 1$ and $-1 < a < 1$.

vi. $\Delta_5 > \Delta_3$

If the following condition is met,

$$\left[1 - \frac{\left(1-a^2\right) g S_S^2}{(2-g) S_R^2 \left(\bar{Y}^2 + S_y^2\right)}\right] > 0. \tag{3.7.2.6}$$

If the above condition is satisfied, this represents that the proposed model II gains maximum protection of the respondent's privacy that Diana and Perri (2010) model.

For the proposed model III, the comparison of the privacy measure are as follows,

vii. $\Delta_6 > \Delta_1$

If $$\left[1 - \frac{S_S^2 \left\{\left(1-g^2\right) + 2g(1-g)\bar{Y}\right\}}{(1-g)^2 \left(S_R^2 + S_S^2\right)\left(\bar{Y}^2 + S_y^2\right)}\right] > 0. \tag{3.7.2.7}$$

Eq. (3.7.2.7) presents that the protection degree of the proposed model III is higher than the additive model presented by Pollock and Bek (1976)

viii. $\Delta_6 > \Delta_2$

If the following condition is met,

$$\left[S_S^2 \left\{\frac{(1-g)^2}{S_R^2 (2-g)} + \frac{\left(g + 2(1-g)\bar{Y}\right)}{S_R^2 (2-g)\left(\bar{Y}^2 + S_y^2\right)}\right\} - 1\right] > 0. \tag{3.7.2.8}$$

Here (3.7.2.8) shows that the privacy protected for the respondents using proposed model III is higher than the multiplicative model.

ix. $\Delta_6 > \Delta_3$

If the following condition is met,

$$\left[(1-g)^2 + \frac{2g(1-g)\bar{Y}}{\left(\bar{Y}^2 + S_y^2\right)}\right] > 0. \quad (3.7.2.9)$$

The condition in (3.7.2.9) presents that the proposed model III attains more privacy protection than Diana and Perri (2010) models.

x. $\Delta_7 > \Delta_1$

If $$\left[1 - S_S^2\left\{\frac{1-g^2\left(1+S_R^2\right)}{S_R^2\left(\bar{Y}^2 + S_y^2\right)}\right\}\right] > 0. \quad (3.7.2.10)$$

This condition satisfies that the proposed model IV gains maximum protection of the respondent's privacy than using additive model.

xi. $\Delta_7 > \Delta_2$

If $g^2 S_S^2\left(1+S_R^2\right) > 0$. (3.7.2.11)

The condition in (3.7.2.11) is always true for $0 < g < 1$.

xii. $\Delta_7 > \Delta_3$

If $$\left[\frac{g S_S^2}{(2-g)\left(\bar{Y}^2 + S_y^2\right)} - 1\right] > 0. \quad (3.7.2.12)$$

It is observed from the above comparisons of proposed model IV provides greater privacy Diana and Perri (2010) *SRR* model.

3.8 DISCUSSION

In this chapter, we have proposed some generalized scrambling randomized response (*SRR*) models for quantitative sensitive responses. The proposed *SRR* models used two random numbers scrambled by multiplying or adding to the sensitive variable in many randomization devices to increase respondent's collaboration. The protection measure for proposed models and some existing models in literature have also been presented. Through these measures, some comparisons between some existing models and proposed models have been discussed. Finally, it is concluded that the proposed models are more protective as compared to some existing models.

CHAPTER 4

PROPOSED ESTIMATORS TO ESTIMATE POPULATION MEAN FOR SINGLE-PHASE SAMPLING BASED ON SCRAMBLED RANDOMIZED RESPONSE MODELS

4.1 INTRODUCTION

In this chapter some estimators using more than one auxiliary variable for the estimation of population mean of sensitive variable y based on scrambled response models. The proposed estimators are presented in the form of observed response z. The expressions of the bias and mean square error have been derived up to first order of approximation.

4.2 PROPOSED ESTIMATOR-1

A generalized regression-cum-exponential estimator has been presented following Cochran (1940) and Upadhaya et al. (2011) as,

$$t_{59}^{(i)} = \left[\bar{z}_j + \omega_1\left(\bar{X}_1 - \bar{x}_1\right)\right]\exp\left[\frac{d_1\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2 + \left(\omega_2 - 1\right)\bar{x}_2}\right],$$

$$\text{for } j = 4,\ G1,\ G2,\ G3,\ G4 \qquad (4.2.1)$$

where $d_1\left(0,-1,+1\right)$ is a constant who generalized the proposed estimator and its different values provides regression estimator, class of regression-cum-exponential ratio and regression-cum-exponential product estimators respectively. ω_1 and ω_2 are assumed to be unknown constants and need to be estimated such that the *MSE* of $t_{59}^{(i)}$ is minimum.

It is observed that regression estimator t_6 in (2.1.16) may be obtained for $d_1 = 0$ and taking $\omega_1 = 0$, we can get a generalized exponential estimator based on single auxiliary variable as described in (2.1.24). However for different choices of these constants, we may get different estimators, e.g. see Table 4.1 provides some examples of different estimators based on two auxiliary variables.

To obtain the bias and mean square error, the proposed estimator $t_{59}^{(i)}$ may be expressed in the form of *e*'s as,

$$t_{59}^{(i)} = \left(\bar{Z}_j\left(1+e_{\bar{z}_j}\right) - \omega_1 \bar{X}_1 e_{\bar{x}_1}\right)\exp\left(\frac{-d_1 e_{\bar{x}_2}}{\omega_2}\left(1+\left(1-\frac{1}{\omega_2}\right)e_{\bar{x}_2}\right)^{-1}\right),$$

or

$$t_{59}^{(i)} = \left(\bar{Z}_j\left(1+e_{\bar{z}_j}\right) - \omega_1 \bar{X}_1 e_{\bar{x}_1}\right)\exp\left(\frac{-d_1 e_{\bar{x}_1}}{\omega_2}\left(1-\left(1-\frac{1}{\omega_2}\right)e_{\bar{x}_1}+\ldots\right)\right). \quad (4.2.2)$$

Table 4.1

Class of Estimators for $t_{59}^{(i)}$

Class of Estimators	ω_1	d_1	ω_2
$t_{59}^{(1)} = \left(\bar{z}_j + \left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right]$	1	1	1
$t_{59}^{(2)} = \left(\bar{z}_j + b_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right]$	$b_{z_j x_1}$	1	1
$t_{59}^{(3)} = \left(\bar{z}_j + \rho_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right]$	$\rho_{z_j x_1}$	1	1
$t_{59}^{(4)} = \left(\bar{z}_j + \left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\bar{X}_2 - \bar{x}_2}{2\left(\bar{X}_2 + \bar{x}_2\right)}\right)\right]$	1	½	2
$t_{59}^{(5)} = \left(\bar{z}_j + b_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\bar{X}_2 - \bar{x}_2}{2\bar{X}_2}\right)\right]$	$b_{z_j x_1}$	½	1
$t_{59}^{(6)} = \left(\bar{z}_j + C_{x1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{2\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right]$	C_{x_1}	2	1
$t_{59}^{(7)} = \left(\bar{z}_j + \rho_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\rho_{z_j x_2}\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right]$	$\rho_{z_j x_1}$	$\rho_{z_j x_2}$	1
$t_{59}^{(8)} = \left(\bar{z}_j + b_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right]$	$b_{z_j x_1}$	1	1
$t_{59}^{(9)} = \left(\bar{z}_j + \rho_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right]$	$\rho_{z_j x_1}$	1	1

The bias and MSE of $t_{59}^{(i)}$, to the first order approximation $O\left(n^{-1}\right)$ may be given as,

$$Bias\left(t_{59}^{(i)}\right) \cong \bar{Z}_j \theta \frac{d_1}{\omega_2}\left[\frac{d_1}{\omega_2} C_{x_2}^2 - \frac{d_1}{\omega_2} C_{z_j} C_{x_2} \rho_{z_j x_2} + \frac{\bar{X}_1}{\bar{Z}_j}\frac{d_1}{\omega_2} C_{x_2} C_{x_1} \rho_{x_1 x_2}\right], \tag{4.2.3}$$

and

$$MSE\left(t_{59}^{(i)}\right) \cong \theta \bar{Z}_j^2 \left[C_{z_j}^2 + \omega^2{}_1 \frac{\bar{X}_1^2}{\bar{Z}_j^2} C^2{}_{x_1} + \frac{d_1^2}{\omega_2} C^2{}_{x_2} - 2\omega_1 \frac{\bar{X}_1}{\bar{Z}_j} C_{z_j} C_{x_1} \rho_{z_j x_1}\right.$$
$$\left. -2\frac{d_1}{\omega_2} C_{z_j} C_{x_2} \rho_{z_j x_2} + 2\omega_1 \frac{d_1}{\omega_2}\frac{\bar{X}_1}{\bar{Z}_j} C_{x_1} C_{x_2} \rho_{x_1 x_2}\right], \tag{4.2.4}$$

and, ω_1 and d_1 attain their optimum values ω_1^{opt} and d_1^{opt} respectively as,

$$\omega_1^{opt} = \frac{\bar{Z}_j C_{z_j}\left(\rho_{z_j x_2}\rho_{x_1 x_2} - \rho_{z_j x_1}\right)}{\bar{X} C_{x_1}\left(1-\rho_{x_1 x_2}^2\right)},$$

and

$$d_1^{opt} = \frac{\omega_2 C_{z_j}\left(\rho_{z_j x_1}\rho_{x_1 x_2} - \rho_{z_j x_2}\right)}{C_{x_2}\left(1-\rho_{x_1 x_2}^2\right)}. \tag{4.2.5}$$

We may get the bias and MSE for $t_{59}^{(i)}$ (i=1,2,3,..,9) using different values of d_1, ω_1 and ω_2 in (4.2.3-4.2.4). Substituting the optimum values in (4.1.4), we may get the minimum value of the $MSE\left(t_{59}^{(i)}\right)$ as,

$$\min MSE\left(t_{59}^{(i)}\right) = \theta \bar{Z}_j^2 C_{z_j}^2 \left[1 + \frac{\left(\rho_{z_j x_2}\rho_{x_1 x_2} - \rho_{z_j x_1}\right)^2}{\left(1-\rho_{x_1 x_2}^2\right)^2} + \frac{\left(\rho_{z_j x_1}\rho_{x_1 x_2} - \rho_{z_j x_2}\right)^2}{\left(1-\rho_{x_1 x_2}^2\right)^2}\right.$$
$$+2\frac{\left(\rho_{z_j x_2}\rho_{x_1 x_2} - \rho_{z_j x_1}\right)\left(\rho_{z_j x_1}\rho_{x_1 x_2} - \rho_{z_j x_2}\right)}{\left(1-\rho_{x_1 x_2}^2\right)^2}$$
$$\left. -2\frac{\rho_{z_j x_2}\left(\rho_{z_j x_1}\rho_{x_1 x_2} - \rho_{z_j x_2}\right)}{\left(1-\rho_{x_1 x_2}^2\right)} - 2\frac{\rho_{z_j x_1}\left(\rho_{z_j x_2}\rho_{x_1 x_2} - \rho_{z_j x_1}\right)}{\left(1-\rho_{x_1 x_2}^2\right)}\right]. \tag{4.2.6}$$

We may get min*MSE*s for $t_{59}^{(i)}$ (i=1,2,3,…,9) using different values of d_1, ω_1 and ω_2 in (4.2.4) (see as given in Table 4.1).

4.3 PROPOSED ESTIMATOR-2

Following Rao (1991) and Upadhaya et al. (2011) a more generalized form of $t_{59}^{(i)}$ is proposed as,

$$t_{60}^{(i)} = \left[\omega_0 \bar{z}_j + \omega_1\left(\bar{X}_1 - \bar{x}_1\right)\right]\exp\left[\frac{d_1\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2 + \left(\omega_2 - 1\right)\bar{x}_2}\right],$$

for j= 4, G1, G2, G3, G4. (4.3.1)

where d_1, ω_1 and ω_2 are the same constants defined for t_{59}^{G}. ω_0 is assumed to be unknown constant and need to be estimated such that the *MSE* of $t_{60}^{(i)}$ is minimum.

It is observed that for $d_1 = 0$ and $\omega_0 = 1$, we obtain the regression estimator t_6 in (2.2.16) and taking $\omega_1 = 0$, we can get a generalized exponential estimator based on single auxiliary variable as described in (2.2.24). However for different choices of these constants, we may get different estimators, e.g. see Table 4.2 provides some examples of different estimators based on two auxiliary variables.

To obtain the bias and mean square error, the proposed estimator $t_{60}^{(i)}$ may be expressed in the form of e's as,

$$t_{60}^{(i)} = \left(\omega_0 \bar{Z}_j\left(1 + e_{\bar{z}_j}\right) - \omega_1 \bar{X}_1 e_{\bar{x}_1}\right)\exp\left(\frac{-d_1 e_{\bar{x}_2}}{\omega_2}\left(1 + \left(1 - \frac{1}{\omega_2}\right)e_{\bar{x}_2}\right)^{-1}\right),$$

or

$$t_{60}^{(i)} = \left(\omega_0 \bar{Z}_j\left(1 + e_{\bar{z}_j}\right) - \omega_1 \bar{X}_1 e_{\bar{x}_1}\right)\exp\left(\frac{-d_1 e_{\bar{x}_1}}{\omega_2}\left(1 - \left(1 - \frac{1}{\omega_2}\right)e_{\bar{x}_1} + \ldots\right)\right).$$

(4.3.2)

Table 4.2
Class of Estimators for $t_{60}^{(i)}$

Class of estimators	ω_0	ω_1	d_1	ω_2
$t_{60}^{(1)} = \left(\omega_0 \bar{z}_j + \left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right]$	ω_0	1	1	1
$t_{60}^{(2)} = \left(\omega_0 \bar{z}_j + b_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right]$	ω_0	$b_{z_j x_1}$	1	1
$t_{60}^{(3)} = \left(\omega_0 \bar{z}_j + \rho_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right]$	ω_0	$\rho_{z_j x_1}$	1	1
$t_{60}^{(4)} = \left(\omega_0 \bar{z}_j + \left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\bar{X}_2 - \bar{x}_2}{2\left(\bar{X}_2 + \bar{x}_2\right)}\right)\right]$	ω_0	1	½	2
$t_{60}^{(5)} = \left(\omega_0 \bar{z}_j + b_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\bar{X}_2 - \bar{x}_2}{2\bar{X}_2}\right)\right]$	ω_0	$b_{z_j x_1}$	½	1
$t_{60}^{(6)} = \left(\omega_0 \bar{z}_j + C_{x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{2\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right]$	ω_0	C_{x_1}	2	1
$t_{60}^{(7)} = \left(\omega_0 \bar{z}_j + \rho_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\rho_{z_j x_2}\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right]$	ω_0	$\rho_{z_j x_1}$	$\rho_{z_j x_2}$	1
$t_{60}^{(8)} = \left(\bar{z}_j + b_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right] = t_{59}^{(8)}$	1	$b_{z_j x_1}$	1	1
$t_{60}^{(9)} = \left(\bar{z}_j + \rho_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right)\right] = t_{59}^{(9)}$	1	$\rho_{z_j x_1}$	1	1

The bias and *MSE* of $t_{60}^{(i)}$, to the first order approximation $O\left(n^{-1}\right)$ may be given as,

$$Bias\left(t_{60}^{(i)}\right) \cong \left(\omega_0 - 1\right)\bar{Z}_j + \omega_0 Bias\left(t_{59}^{(i)}\right), \tag{4.3.3}$$

and

$$MSE\left(t_{60}^{(i)}\right) \cong \left(\omega_0 - 1\right)^2 \bar{Z}_j^2 + \omega_0^2 MSE\left(t_{59}^{(i)}\right) + 2\omega_0\left(\omega_0 - 1\right)\bar{Z}_j Bias\left(t_{59}^{(i)}\right). \tag{4.3.4}$$

and the $MSE\left(t_{60}^{(i)}\right)$ is minimum for the optimum value of ω_0 as,

$$\omega_0^{opt} = \frac{\bar{Z}_j\left[\bar{Z}_j + Bias\left(t_{59}^{(i)}\right)\right]}{\left[\bar{Z}_j^2 + MSE\left(t_{59}^{(i)}\right) + 2\bar{Z}_j Bias\left(t_{59}^{(i)}\right)\right]}. \tag{4.3.5}$$

Substituting the optimum value of ω_0 in (4.3.4), we may get the minimum value of the $MSE\left(t_{60}^{(i)}\right)$ as,

$$\min MSE\left(t_{60}^{(i)}\right) = \bar{Z}_j^2\left(1 - \frac{\left[1 + \frac{Bias\left(t_{59}^{(i)}\right)}{\bar{Z}_j}\right]}{\left[1 + \frac{MSE\left(t_{59}^{(i)}\right)}{\bar{Z}_j^2} + 2\frac{Bias\left(t_{59}^{(i)}\right)}{\bar{Z}_j}\right]}\right). \tag{4.3.6}$$

From (4.3.4)-(4.3.6) it is clear that the $MSE\left(t_{60}^{(i)}\right)$ and $\min MSE\left(t_{60}^{(i)}\right)$ are the functions of $Bias\left(t_{59}^{(i)}\right)$ and $MSE\left(t_{59}^{(i)}\right)$. Also, one can get (4.3.4)-(4.3.6) if $Bias\left(t_{59}^{(i)}\right)$ and $MSE\left(t_{59}^{(i)}\right)$ are known.

4.4 PROPOSED ESTIMATOR-3

Following Gupta et al. (2012) a generalized regression-cum-ratio estimator has been presented as follows:

$$t_{61}^{(i)} = \left[\bar{z}_j + \omega_3\left(\bar{X}_1 - \bar{x}_1\right)\right]\left[\frac{\bar{X}_2}{\bar{x}_2}\right]^{k_1}, \text{ for } j = 4, \text{G1, G2, G3, G4}. \tag{4.4.1}$$

where ω_3 is assumed to be unknown constant and need to be estimated such that the *MSE* of $t_{61}^{(i)}$ is minimum. Also, $k_1(0,-1,+1)$ is a suitable constant whose values yields different ratio and product estimators. From the proposed estimator given in (4.4.1), it is noticed that regression estimator t_6 in (2.1.16) may be obtained for $k_1 = 0$, and for $\omega_3 = 0$ and $k_1 = 1$ we obtain ratio estimator based on single auxiliary variable in (2.1.1). However for different choices of these constants, we may get different estimators, e.g. see Table 4.3 provides some example of different estimators based on two auxiliary variables.

Table 4.3

Class of Estimators for $t_{61}^{(i)}$

Class of estimators	ω_3	k_1
$t_{61}^{(1)} = \left[\bar{z}_j + \left(\bar{X}_1 - \bar{x}_1\right)\right]\left[\frac{\bar{X}_2}{\bar{x}_2}\right]$	1	1
$t_{61}^{(2)} = \left[\bar{z}_j + \rho_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right]\left[\frac{\bar{X}_2}{\bar{x}_2}\right]$	$\rho_{z_j x_1}$	1
$t_{61}^{(3)} = \left[\bar{z}_j + \left(\bar{X}_1 - \bar{x}_1\right)\right]\left[\frac{\bar{x}_2}{\bar{X}_2}\right]$	1	-1
$t_{61}^{(4)} = \left[\bar{z}_j + b_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right]\left[\frac{\bar{x}_2}{\bar{X}_2}\right]$	$b_{z_j x_1}$	-1
$t_{61}^{(5)} = \left[\bar{z}_j + \rho_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1\right)\right]\left[\frac{\bar{x}_2}{\bar{X}_2}\right]$	$\rho_{z_j x_1}$	-1

To obtain the bias and mean square error, the proposed estimator $t_{61}^{(i)}$ may be expressed in the form of e's as,

$$t_{61}^{(i)} = \left[\omega_3 \bar{Z}_j \left(1 + e_{\bar{z}}\right) - \omega_4 \bar{X}_1 e_{\bar{x}_1}\right]\left(1 + e_{\bar{x}_2}\right)^{-k_1}, \qquad (4.4.2)$$

Solving (4.4.2), we may have,

$$t_{61}^{(i)} - \bar{Z}_j = \bar{Z}_j \Bigg[1 + e_{\bar{z}_j} - k_1 e_{\bar{x}_2} - k_1 e_{\bar{z}_j} e_{\bar{x}_2} + \frac{k_1\left(k_1 + 1\right)}{2} e_{\bar{x}_2}^2 - \omega_3 \frac{\bar{X}_1}{\bar{Z}_j}\left(e_{\bar{x}_1} - k_1 e_{\bar{x}_1} e_{\bar{x}_2}\right)\Bigg]. \qquad (4.4.3)$$

The expressions of the bias and *MSE* of $t_{61}^{(i)}$ are obtained as,

$$Bias\left(t_{61}^{(i)}\right)=\bar{Z}_j\theta k_1\left[\frac{(k_1+1)}{2}C_{x2}-C_{z_j}\rho_{z_jx2}+\omega_3\frac{\bar{X}_1}{\bar{Z}_j}C_{x1}\rho_{x1x2}\right], \quad (4.4.4)$$

and

$$MSE\left(t_{61}^{(i)}\right)=\bar{Z}_j^2\theta\left[C_{z_j}^2+k_1^2C_{x_2}^2+\omega_4^2\frac{\bar{X}_1^{\,2}}{\bar{Z}_j^2}C_{x_1}^2\right.$$
$$\left.+2\omega_3\omega_4\frac{\bar{X}_1}{\bar{Z}_j}C_{x_1}\left(C_{z_j}\rho_{z_jx_1}-k_1C_{x_2}\rho_{x_1x_2}\right)-2k_1C_{z_j}C_{x_2}\rho_{z_jx_2}\right]. \quad (4.4.5)$$

The MSE of $t_{61}^{(i)}$ is minimum for the optimum values of ω_3 and k_1 are as follows:

$$\omega_3^{opt}=\frac{\bar{Z}_jC_{z_j}\left(\rho_{z_jx_2}\rho_{x_1x_2}-\rho_{z_jx_1}\right)}{\bar{X}_1C_{x_1}\left(1-\rho_{x_1x_2}^2\right)},$$

and

$$k_1^{opt}=\frac{C_{z_j}\left(\rho_{z_jx_1}\rho_{x_1x_2}-\rho_{z_jx_2}\right)}{C_{x_2}\left(1-\rho_{x_1x_2}^2\right)}. \quad (4.4.6)$$

From (4.4.4-4.4.5), we may get the bias and MSE for the estimator $t_{61}^{(i)}$ (i=1,2,…,5) using different values of ω_3 and k_1 as given in Table 4.3. Using (4.4.6), the minimum MSE obtained presented as,

$$\min MSE\left(t_{61}^{(i)}\right)=\theta\bar{Z}_j^2C^2{}_Z\left[1+\frac{\left(\rho_{z_jx_2}\rho_{x_1x_2}-\rho_{z_jx_1}\right)^2}{\left(1-\rho_{x_1x_2}^2\right)^2}+\frac{\left(\rho_{z_jx_1}\rho_{x_1x_2}-\rho_{z_jx_2}\right)^2}{\left(1-\rho_{x_1x_2}^2\right)^2}\right.$$
$$+2\frac{\left(\rho_{z_jx_2}\rho_{x_1x_2}-\rho_{z_jx_1}\right)\left(\rho_{z_jx_1}\rho_{x_1x_2}-\rho_{z_jx_2}\right)}{\left(1-\rho_{x_1x_2}^2\right)^2}$$
$$\left.-2\frac{\rho_{z_jx_2}\left(\rho_{z_jx_1}\rho_{x_1x_2}-\rho_{z_jx_2}\right)}{\left(1-\rho_{x_1x_2}^2\right)}-2\frac{\rho_{z_jx_1}\left(\rho_{z_jx_2}\rho_{x_1x_2}-\rho_{z_jx_1}\right)}{\left(1-\rho_{x_1x_2}^2\right)}\right]. \quad (4.4.7)$$

We may get min *MSE*s for $t_{61}^{(i)}$ (i=1,2,…,5) using different values of ω_3 and k_1 in (4.4.5) (e.g. as given in Table 4.3).

4.5 PROPOSED ESTIMATOR-4

Another generalized regression-cum-ratio estimator has been proposed by following Searls (1964) and Gupta (2012) is given by,

$$t_{62}^{(i)} = \left[\omega_4 \bar{z}_j + \omega_3\left(\bar{X}_1 - \bar{x}_1\right)\right]\left[\frac{\bar{X}_2}{\bar{x}_2}\right]^{k_1}, \text{ for } j = 4, \text{G1, G2, G3, G4} \quad (4.5.1)$$

where ω_4 is assumed to be unknown constant and need to be estimated such that the *MSE* of $t_{62}^{(i)}$ is minimum. k_1 and ω_3 are the same constants defined for $t_{61}^{(i)}$. From the proposed estimator given in (4.1.20), it is noticed that the regression estimator t_6 in (2.2.16) may be obtained for $\omega_4 = 1$ and $k = 0$, and, for $\omega_3 = 0$, $\omega_4 = 1$ and $k = 1$ we obtain ratio estimator based on single auxiliary variable in (2.2.1). However for different choices of these constants, we may get different estimators, e.g. see Table 4.4 provides some example of different estimators based on two auxiliary variables.

To obtain the bias and mean square error, the proposed estimator $t_{62}^{(i)}$ may be expressed in the form of e's as,

$$t_{62}^{(i)} = \left[\omega_4 \bar{Z}_j\left(1 + e_{\bar{z}}\right) - \omega_3 \bar{X}_1 e_{\bar{x}_1}\right]\left(1 + e_{\bar{x}_2}\right)^{-k_1}. \quad (4.5.2)$$

Table 4.4

Class of Estimators for $t_{62}^{(i)}$

Class of estimators	ω_3	ω_4	k_1
$t_{62}^{(1)}=\left[\omega_4\bar{z}_j+\left(\bar{X}_1-\bar{x}_1\right)\right]\left[\dfrac{\bar{X}_2}{\bar{x}_2}\right]$	1	ω_4	1
$t_{62}^{(2)}=\left[\omega_4\bar{z}_j+b_{z_jx_1}\left(\bar{X}_1-\bar{x}_1\right)\right]\left[\dfrac{\bar{X}_2}{\bar{x}_2}\right]$	$b_{z_jx_1}$	ω_4	1
$t_{62}^{(3)}=\left[\omega_4\bar{z}_j+\rho_{z_jx_1}\left(\bar{X}_1-\bar{x}_1\right)\right]\left[\dfrac{\bar{X}_2}{\bar{x}_2}\right]$	$\rho_{z_jx_1}$	ω_4	1
$t_{62}^{(4)}=\left[\omega_4\bar{z}_j+\left(\bar{X}_1-\bar{x}_1\right)\right]\left[\dfrac{\bar{x}_2}{\bar{X}_2}\right]$	1	ω_4	-1
$t_{62}^{(5)}=\left[\omega_4\bar{z}_j+b_{z_jx_1}\left(\bar{X}_1-\bar{x}_1\right)\right]\left[\dfrac{\bar{x}_2}{\bar{X}_2}\right]$	$b_{z_jx_1}$	ω_4	-1
$t_{62}^{G6}=\left[\omega_4\bar{z}_j+\rho_{z_jx_1}\left(\bar{X}_1-\bar{x}_1\right)\right]\left[\dfrac{\bar{x}_2}{\bar{X}_2}\right]$	$\rho_{z_jx_1}$	ω_4	-1
$t_{62}^{G7}=\left[\bar{z}_j+b_{z_jx_1}\left(\bar{X}_1-\bar{x}_1\right)\right]\left[\dfrac{\bar{X}_2}{\bar{x}_2}\right]$	$b_{z_jx_1}$	1	1
$t_{62}^{G8}=\left[\bar{z}_j+b_{z_jx_1}\left(\bar{X}_1-\bar{x}_1\right)\right]\left[\dfrac{\bar{x}_2}{\bar{X}_2}\right]$	$b_{z_jx_1}$	1	-1

Solving (4.5.2), we may have,

$$t_{62}^{(i)}-\bar{Z}_j=\left(\omega_4-1\right)\bar{Z}_j+\bar{Z}_j\left[\omega_4\left(e_{\bar{z}_j}-k_1e_{\bar{x}_2}-k_1e_{\bar{z}_j}e_{\bar{x}_2}+\frac{k_1\left(k_1+1\right)}{2}e_{\bar{x}_2}^2\right)\right.$$
$$\left.-\omega_3\frac{\bar{X}_1}{\bar{Z}_j}\left(e_{\bar{x}_1}-k_1e_{\bar{x}_1}e_{\bar{x}_2}\right)\right]. \qquad (4.5.3)$$

The expressions of the bias and *MSE* of $t_{62}^{(i)}$ are obtained as,

$$Bias\left(t_{62}^{(i)}\right)\cong\left(\omega_0-1\right)\bar{Z}_j+\omega_0 Bias\left(t_{61}^{(i)}\right), \qquad (4.5.4)$$

and

$$MSE\left(t_{62}^{(i)}\right) \cong (\omega_4 - 1)^2 \bar{Z}_j^2 + \omega_4^2 MSE\left(t_{61}^{(i)}\right) + 2\omega_4(\omega_4 - 1)\bar{Z}_j Bias\left(t_{61}^{(i)}\right), \tag{4.5.5}$$

and the $MSE\left(t_{62}^{(i)}\right)$ is minimum for the optimum value ω_4 as ω_4^{opt} is given by,

$$\omega_4^{opt} = \frac{\bar{Z}_j\left[\bar{Z}_j + Bias\left(t_{61}^{(i)}\right)\right]}{\left[\bar{Z}_j^2 + MSE\left(t_{61}^{(i)}\right) + \bar{Z}_j Bias\left(t_{61}^{(i)}\right)\right]}. \tag{4.5.6}$$

Substituting the optimum value of ω_4 in (4.5.5), we may get the minimum value of the $MSE\left(t_{62}^{(i)}\right)$ as,

$$\min MSE\left(t_{62}^{(i)}\right) = \bar{Z}^2\left(1 - \frac{\left[1 + \frac{Bias\left(t_{61}^{(i)}\right)}{\bar{Z}}\right]}{\left[1 + \frac{MSE\left(t_{61}^{(i)}\right)}{\bar{Z}^2} + 2\frac{Bias\left(t_{61}^{(i)}\right)}{\bar{Z}}\right]}\right). \tag{4.5.7}$$

From (4.5.5)-(4.5.7), it is clearly observed that $MSE\left(t_{62}^{(i)}\right)$ and $\min MSE\left(t_{62}^{(i)}\right)$ are the functions of $Bias\left(t_{61}^{(i)}\right)$ and $MSE\left(t_{61}^{(i)}\right)$.

4.6 PROPOSED ESTIMATOR-5

Another estimator proposed using two auxiliary variables when there exists exponential relation between sensitive study variable and auxiliary variables. The proposed estimator presented in the form of observed response z as,

$$t_{63} = \bar{z}_j \exp\left[\frac{b'\left(\bar{X}_1 - \bar{x}_1\right)}{\bar{X}_1 + (c' - 1)\bar{x}_1}\right] \exp\left[\frac{b''\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2 + (c'' - 1)\bar{x}_2}\right].$$

for j= 4, G1, G2, G3, G4. (4.6.1)

Rewrite (4.6.1) in the form of *e*'s as,

$$t_{63} = \bar{Z}_j\left(1+e_{\bar{z}_j}\right)\exp\left[\frac{-b'e_{\bar{x}_1}}{c'\left(1+e_{\bar{x}_1}\right)}\left(1+\frac{e_{\bar{x}_1}}{c'\left(1+e_{\bar{x}_1}\right)}\right)^{-1}\right]$$
$$\exp\left[\frac{-b''e_{\bar{x}_2}}{c''\left(1+e_{\bar{x}_2}\right)}\left(1+\frac{e_{\bar{x}_2}}{c''\left(1+e_{\bar{x}_2}\right)}\right)^{-1}\right]. \qquad (4.6.2)$$

Solving (4.6.2), we may have,

$$t_{63} - \bar{Z}_j = \bar{Z}_j\left[\begin{array}{l} e_{\bar{z}_j} - d'e_{\bar{x}_1} - d'e_{\bar{x}_1}e_{\bar{z}_j} + d'^2e_{\bar{x}_1}^2 - d''e_{\bar{x}_2} \\ \qquad -d''e_{\bar{x}_2}e_{\bar{z}_j} + d''^2e_{\bar{x}_2}^2 + d'd''e_{\bar{x}_1}e_{\bar{x}_2} \end{array}\right], \qquad (4.6.3)$$

where $d' = b'/c'$ and $d'' = b''/c''$.

The expressions of the bias and *MSE* of t_{63} are obtained as,

$$Bias\left(t_{63}\right) = \bar{Z}_j\theta\left[d'C_{x_1}^2 + d''C_{x_2}^2 - d'C_{z_jx_1} + d'd''C_{x_1x_2} - d''C_{z_jx_2}\right], \qquad (4.6.4)$$

and

$$MSE\left(t_{63}\right) = \bar{Z}_j^2\theta\left[\begin{array}{l} C_{z_j}^2 + d'^2C_{x_1}^2 + d''^2C_{x_2}^2 \\ \quad + 2d'd''C_{x_1x_2} - 2d'C_{z_jx_1} - 2d''C_{z_jx_2} \end{array}\right]. \qquad (4.6.5)$$

The optimum values we attain for d' and d'' are respectively as,

$$d'^{opt} = \frac{C_{z_j}\left(\rho_{z_jx_1} - \rho_{z_jx_2}\rho_{x_1x_2}\right)}{C_{x_1}\left(1-\rho_{x_1x_2}^2\right)}, \qquad (4.6.6)$$

and

$$d'' = \frac{C_{z_j}\left(\rho_{z_jx_2} - \rho_{z_jx_1}\rho_{x_1x_2}\right)}{C_{x_2}\left(1-\rho_{x_1x_2}^2\right)}. \qquad (4.6.7)$$

Using (4.6.6-4.6.7), the minimum *MSE* obtained presented as,

$$\min MSE\left(t_{63}\right)=\theta\bar{Z}_j^2 C^2{}_Z\left[1+\frac{\left(\rho_{z_j x_1}-\rho_{z_j x_2}\rho_{x_1 x_2}\right)^2}{\left(1-\rho_{x_1 x_2}^2\right)^2}+\frac{\left(\rho_{z_j x_2}-\rho_{z_j x_1}\rho_{x_1 x_2}\right)^2}{\left(1-\rho_{x_1 x_2}^2\right)^2}\right.$$

$$+2\frac{\left(\rho_{z_j x_1}-\rho_{z_j x_2}\rho_{x_1 x_2}\right)\left(\rho_{z_j x_2}-\rho_{z_j x_1}\rho_{x_1 x_2}\right)}{\left(1-\rho_{x_1 x_2}^2\right)^2}$$

$$\left.-2\frac{\rho_{z_j x_2}\left(\rho_{z_j x_2}-\rho_{z_j x_1}\rho_{x_1 x_2}\right)}{\left(1-\rho_{x_1 x_2}^2\right)}-2\frac{\rho_{z_j x_1}\left(\rho_{z_j x_1}-\rho_{z_j x_2}\rho_{x_1 x_2}\right)}{\left(1-\rho_{x_1 x_2}^2\right)}\right]. \quad (4.6.8)$$

4.7 PROPOSED ESTIMATOR-6

A new chain ratio in regression-cum-exponential estimator has been proposed is presented as:

$$t_{64}^{(i)}=\left[\bar{z}_j+\omega_5\left(\bar{X}_1-\bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}\right)\right]\exp\left[\frac{d_2\left(\bar{X}_2-\bar{x}_2\right)}{\bar{X}_2+\left(c_1-1\right)\bar{x}_2}\right],$$

for j=4, G1, G2, G3, G4. (4.7.1)

where $d_2\left(0,-1,+1\right)$ is a constant who generalized the proposed estimator and its different values provides ratio in regression estimator, class of regression-cum-exponential ratio and regression-cum-exponential product estimators respectively. ω_5 and c_1 are assumed to be unknown constants and need to be estimated such that the *MSE* of $t_{64}^{(i)}$ is minimum. Table 4.5 provides some example of different estimator for different choices of ω_5, d_2 and c_1.

Rewrite (4.7.1) in the form of *e*'s as,

$$t_{64}^{(i)}=\left[\bar{Z}_j\left(1+e_{\bar{z}_j}\right)-\omega_5\bar{X}_1\left(e_{\bar{x}_1}-e_{\bar{x}_2}\right)\right]\exp\left[\frac{-d_2 e_{\bar{x}_2}}{c_1\left(1+e_{\bar{x}_2}\right)}\left(1-\frac{e_{\bar{x}_2}}{c_1\left(1+e_{\bar{x}_2}\right)}\right)^{-1}\right],$$

(4.7.2)

Solving (4.7.2), we may have,

$$t_{64}^{(i)} - \bar{Z}_j = \bar{Z}_j \left[e_{\bar{z}_j} - \frac{d_2}{c_1}\left(e_{\bar{x}_2} - e_{\bar{x}_2}^2 - e_{\bar{z}_j} e_{\bar{x}_2}\right) + \frac{d_2^3}{2c_1^2} e_{\bar{x}_2}^2 \right.$$

$$\left. -\omega_5 \frac{\bar{X}_1}{\bar{Z}_j}\left\{ e_{\bar{x}_1} - e_{\bar{x}_2} - \frac{d_2}{c_1}\left(e_{\bar{x}_1} e_{\bar{x}_2} + e_{\bar{x}_1}^2\right)\right\}\right]. \qquad (4.7.3)$$

Table 4.5

Class of Estimators for $t_{64}^{(i)}$

Class of Estimators	ω_5	d_2	c_1
$t_{64}^{(1)} = \left[\bar{z}_j + \left(\bar{X}_1 - \bar{x}_1 \frac{\bar{X}_2}{\bar{x}_2}\right)\right] \exp\left[\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right]$	1	1	1
$t_{64}^{(2)} = \left[\bar{z}_j + b_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1 \frac{\bar{X}_2}{\bar{x}_2}\right)\right] \exp\left[\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right]$	$b_{z_j x_1}$	1	1
$t_{64}^{(3)} = \left[\bar{z}_j + \rho_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1 \frac{\bar{X}_2}{\bar{x}_2}\right)\right] \exp\left[\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right]$	$\rho_{z_j x_1}$	1	1
$t_{64}^{(4)} = \left[\bar{z}_j + \left(\bar{X}_1 - \bar{x}_1 \frac{\bar{X}_2}{\bar{x}_2}\right)\right] \exp\left[\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{2\left(\bar{X}_2 + \bar{x}_2\right)}\right]$	1	½	2
$t_{64}^{(5)} = \left[\bar{z}_j + b_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1 \frac{\bar{X}_2}{\bar{x}_2}\right)\right] \exp\left[\frac{\left(\bar{X}_2 - \bar{x}_2\right)}{2\left(\bar{X}_2 + \bar{x}_2\right)}\right]$	$b_{z_j x_1}$	½	1
$t_{64}^{(6)} = \left[\bar{z}_j + C_{x_1}\left(\bar{X}_1 - \bar{x}_1 \frac{\bar{X}_2}{\bar{x}_2}\right)\right] \exp\left[\frac{2\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2}\right]$	C_{x_1}	2	1

The expressions of the bias and *MSE* of $t_{64}^{(i)}$ are obtained as,

$$Bias\left(t_{64}^{(i)}\right) = \bar{Z}_j \theta \frac{d_2}{c_1}\left[C_{x_2}^2 - C_{z_j x_2} + \frac{d_2}{2c_1} C_{x_2}^2 - \omega_5 \frac{\bar{X}_1}{\bar{Z}_j}\left(C_{x_1 x_2} - C_{x_2}^2\right)\right], \quad (4.7.4)$$

and

$$MSE\left(t_{64}^{(i)}\right) = \bar{Z}_j^2 \theta \left[C_{z_j}^2 + \frac{d_2^2}{c_1^2} C_{x_2}^2 + \omega_5^2 \frac{\bar{X}_1^2}{\bar{Z}_j^2}\left(C_{x_1}^2 + C_{x_2}^2 - 2C_{x_1 x_2}\right) \right.$$

$$\left. -2\frac{d_2}{c_1} C_{z_j x_2} - 2\omega_5 \frac{\bar{X}_1}{\bar{Z}_j}\left(C_{z_j x_1} - C_{z_j x_2}\right) + 2\omega_5 \frac{\bar{X}_1}{\bar{Z}_j}\frac{d_2}{c_1}\left(C_{x_1 x_2} - C_{x_2}^2\right)\right]. \quad (4.7.5)$$

The MSE of $t_{64}^{(i)}$ is minimum for the optimum values of ω_5 and d_2 are as follows:

$$\omega_5^{opt} = \frac{\bar{Z}_j C_{z_j}\left(C_{x_1}\rho_{z_j x_1} - C_{x_2}\rho_{z_j x_2}\right)}{\bar{X}_1 C_{x_1}\left(C_{x_1} - 2C_{x_2}\rho_{x_1 x_2}\right)} = \frac{\bar{Z}_j C_{z_j} K_1}{\bar{X}_1 C_{x_1} K_2},$$

and

$$d_2^{opt} = \frac{c_1 C_{z_j}\left(2C_{x_1}\rho_{z_j x_1} - C_{x_2}\rho_{z_j x_2}\right)}{C_{x_1 x_2}} = \frac{c_1 C_{z_j} K_3}{C_{x_1 x_2}}. \quad (4.7.6)$$

We may get the bias and MSE for the estimator $t_{64}^{(i)}$ $(i = 1, 2, \ldots, 6)$ using different values of ω_5 and d_2 as presented in Table 4.5. Substituting the optimum values of ω_5^{opt} and $d_2{}^{opt}$ in (4.7.5), the expression of $\min MSE\left(t_{64}^{(i)}\right)$ we may obtain is given by,

$$\min MSE\left(t_{64}^{(i)}\right) = \bar{Z}_j^2 \theta C_{z_j}^2 \left[1 + \frac{1}{C_{x_1}^2}\left\{\frac{K_3^2}{\rho_{x_1 x_2}^2} + \frac{K_1^2}{K_2^2}\left(C_{x_1}^2 + C_{x_2}^2 - 2C_{x_1 x_2}\right)\right.\right.$$
$$-2\frac{K_3}{\rho_{x_1 x_2}} C_{x_1}\rho_{z_j x_2} - 2C_{x_1}\frac{K_1}{K_2}\left(C_{x_1}\rho_{z_j x_1} - C_{x_2}\rho_{z_j x_2}\right)$$
$$\left.\left. + 2\frac{K_1}{K_3}\frac{K_3}{\rho_{x_1 x_2}}\left(C_{x_1}\rho_{x_1 x_2} - C_{x_2}\right)\right\}\right]. \quad (4.7.8)$$

We may get the minMSE's for $t_{64}^{(i)}$ $(i = 1, 2, \ldots, 6)$ using different values of ω_5 and d_2 as given in Table 4.5.

4.8 PROPOSED ESTIMATOR-7

Another new chain ratio in regression-cum-ratio estimator has been proposed as follows:

$$t_{65}^{(i)} = \left[\bar{z}_j + \omega_6\left(\bar{X}_1 - \bar{x}_1\frac{\bar{X}_1}{\bar{x}_2}\right)\right]\left[\frac{\bar{X}_2}{\bar{x}_2}\right]^{k_2},$$

for j=4, G1, G2, G3, G4. (4.8.1)

where $k_2\left(0,-1,+1\right)$ is a constant who generalized the proposed estimator and its different values provides ratio-cum-regression estimators. Also, ω_6 is assumed to be unknown constant and need to be estimated such that the *MSE* of $t_{65}^{(i)}$ is minimum. In Table 4.6, different estimators are presented for different choices of ω_6 and k_2.

To obtain the bias and mean square error, the proposed estimator $t_{65}^{(i)}$ may be expressed in the form of *e*'s as,

$$t_{65}^{(i)}=\left(\bar{Z}_j\left(1+e_{\bar{z}_j}\right)-\omega_6\bar{X}_1\left(e_{\bar{x}_1}-e_{\bar{x}_2}-e_{\bar{x}_1}e_{\bar{x}_2}\right)\right)\left(1+e_{\bar{x}_2}\right)^{-k_2},$$

or

$$t_{65}^{(i)}=\left(\bar{Z}_j\left(1+e_{\bar{z}_j}\right)-\omega_6\bar{X}_1\left(e_{\bar{x}_1}-e_{\bar{x}_2}-e_{\bar{x}_1}e_{\bar{x}_2}\right)\right)\left(1-k_2e_{\bar{x}_2}+\frac{k_2\left(k_2+1\right)}{2}e_{\bar{x}_2}^2\right),$$

or

$$t_{65}^{(i)}=\bar{Z}_j\left[1+e_{\bar{z}j}-k_2e_{\bar{x}_2}+\frac{k_2\left(k_2+1\right)}{2}e_{\bar{x}_2}^2-k_2e_{\bar{z}_j}e_{\bar{x}_2}\right.$$
$$\left.-\frac{\bar{X}_1}{\bar{Z}_j}\omega_6\left(e_{\bar{x}_1}-e_{\bar{x}_2}-e_{\bar{x}_1}e_{\bar{x}_2}-k_2e_{\bar{x}_2}e_{\bar{x}_1}+k_2e_{\bar{x}_2}^2\right)\right]. \qquad (4.8.2)$$

Table 4.6

Class of Estimators for $t_{65}^{(i)}$

Class of estimators	ω_6	k_2
$t_{65}^{(1)} = \left[\bar{z}_j + \omega_6\left(\bar{X}_1 - \bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}\right)\right]\left[\frac{\bar{X}_2}{\bar{x}_2}\right]$	ω_6	1
$t_{65}^{(2)} = \left[\bar{z}_j + \omega_6\left(\bar{X}_1 - \bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}\right)\right]\left[\frac{\bar{x}_2}{\bar{X}_2}\right]$	ω_6	-1
$t_{65}^{(3)} = \left[\bar{z}_j + \omega_6\left(\bar{X}_1 - \bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}\right)\right]$	ω_6	0
$t_{65}^{(4)} = \left[\bar{z}_j + \left(\bar{X}_1 - \bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}\right)\right]\left[\frac{\bar{X}_2}{\bar{x}_2}\right]$	1	1
$t_{65}^{(5)} = \left[\bar{z}_j + b_{zx1}\left(\bar{X}_1 - \bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}\right)\right]\left[\frac{\bar{X}_2}{\bar{x}_2}\right]$	b_{zx1}	½
$t_{65}^{(6)} = \left[\bar{z}_j + \rho_{zx1}\left(\bar{X}_1 - \bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}\right)\right]\left[\frac{\bar{X}_2}{\bar{x}_2}\right]$	ρ_{zx1}	2

The bias and *MSE* of $t_{65}^{(i)}$, to the first order approximation $O\left(n^{-1}\right)$ may be given as,

$$Bias\left(t_{65}^{(i)}\right) \cong \theta\bar{Z}_j\left[\frac{k_2\left(k_2+1\right)}{2}C_{x_2}^2 - k_2C_{z_jx_2} - \frac{\bar{X}_1}{\bar{Z}_j}\omega_6\left(k_2C_{x_2}^2 - \left(k_2+1\right)C_{x_1x_2}\right)\right], \tag{4.8.3}$$

and

$$MSE\left(t_{65}^{(i)}\right) \cong \theta\bar{Z}_j^2\left[\left(C_{z_j}^2 + k_2^2C_{x_2}^2 - 2k_2C_{z_jx_2}\right) + \frac{\bar{X}_1^2}{\bar{Z}_j^2}\omega_6^2\left(C_{x_1}^2 + C_{x_2}^2 - 2C_{x_1x_2}\right)\right.$$
$$\left. -2\frac{\bar{X}_1}{\bar{Z}_j}\omega_6\left(C_{z_jx_2} - C_{z_jx_2} - k_2C_{x_1x_2} + k_2C_{x_2}^2\right)\right]. \tag{4.8.4}$$

The constants ω_6 and k_2 attain their optimum values ω_6^{opt} and k_2^{opt} respectively as,

$$\omega_6^{opt} = \frac{\bar{Z}_j C_{z_j} \left(\rho_{z_j x_1} + \rho_{z_j x_2}\rho_{x_1 x_2}\right)}{\bar{X}_1 C_{x_1}\left(1-\rho_{x_1 x_2}^2\right)}$$

and

$$k_2^{opt} = \frac{1}{C_{X_2}}\left(C_{z_j}\rho_{z_j x_2} + \frac{\bar{X}_1}{\bar{Z}_j}\left(C_{x_1}\rho_{x_1 x_2} - C_{x_2}\right)\right) \quad (4.8.5)$$

We may get the bias and *MSE* for the estimator $t_{65}^{(i)}$ $(i=1,2,\ldots,6)$ using different values of ω_6 and k_2 given in Table 4.6. Substituting the optimum values in (4.8.4), we may get the minimum value of the $MSE\left(t_{65}^{(i)}\right)$ as,

$$\min MSE\left(t_{65}^{(i)}\right) = \theta \bar{Z}_j^2 C_{z_j}^2 \left[\left(1-\rho_{z_j x_2}^2\right) - \frac{\left(\rho_{z_j x_1} + \rho_{z_j x_2}\rho_{x_1 x_2}\right)^2}{\left(1-\rho_{x_1 x_2}^2\right)}\right]. \quad (4.8.6)$$

From (4.8.6), we may get the min*MSE*'s for $t_{65}^{(i)}$ $(i=1,2,\ldots,6)$ using different values of ω_6 and k_2 as given in Table 4.6.

4.9 PROPOSED ESTIMATOR-8

Another generalized exponential-cum-exponential estimator has been proposed as,

$$t_{66}^{(i)} = \bar{z}_j \exp\left[\frac{d_3\left(\bar{X}_1 - \bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}\right)}{\bar{X}_1 + \left(\beta_1 - 1\right)\bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}}\right]\exp\left[\frac{d_4\left(\bar{X}_2 - \bar{x}_2\right)}{\bar{X}_2 + \left(\beta_2 - 1\right)\bar{x}_2}\right],$$

for j=4, G1, G2, G3, G4 (4.9.1)

where d_3 and d_4 are constants who generalized the proposed estimator and their different values provides exponential estimators. β_1 and β_2 are assumed to be unknown constants and need to be estimated such that the *MSE* of $t_{66}^{(i)}$ is minimum. For different choices of these constants, we may get different estimators, e.g. see Table 4.7 provides some examples of different estimators based on two auxiliary variables.

To obtain the bias and mean square error, the proposed estimator $t_{66}^{(i)}$ may be expressed in the form of e’s as,

$$t_{66}^{(i)} = \bar{Z}_j\left(1+e_{\bar{z}_j}\right)\exp\left(\frac{d_3\left(e_{\bar{x}_1}-e_{\bar{x}_2}\right)}{\beta_1}\left(1+\left(e_{\bar{x}_1}-e_{\bar{x}_2}\right)\right)^{-1}\right)\exp\left(\frac{-d_4 e_{\bar{x}_2}}{\beta_2}\left(1+e_{\bar{x}_2}\right)^{-1}\right),$$

or

$$t_{66}^{(i)} = \bar{Z}_j\left(1+e_{\bar{z}_j}\right)\left(1+\nu_3\left(e_{\bar{x}_1}-e_{\bar{x}_2}\right)+\frac{1}{2}\nu_3^2\left(e_{\bar{x}_1}-e_{\bar{x}_2}\right)^2\right)\left(1-\nu_4 e_{\bar{x}_2}+\frac{1}{2}\nu_4^2 e_{\bar{x}_2}^2\right),$$

or

$$t_{66}^{(i)} - \bar{Z}_j = \bar{Z}_j\Big[e_{\bar{z}_j}+\nu_3\left(e_{\bar{x}_1}-e_{\bar{x}_2}\right)+\nu_3\left(e_{\bar{x}_1}e_{\bar{z}_j}-e_{\bar{x}_2}e_{\bar{z}_j}\right)+\frac{1}{2}\nu_3^2\left(e_{\bar{x}_1}-e_{\bar{x}_2}\right)^2$$
$$-\nu_4 e_{\bar{x}_2}-\nu_4 e_{\bar{z}_j}e_{\bar{x}_2}-\nu_3\nu_4\left(e_{\bar{x}_2}e_{\bar{x}_1}-e_{\bar{x}_2}^2\right)+\frac{1}{2}\nu_4^2 e_{\bar{x}_2}^2\Big], \qquad (4.9.2)$$

where $\nu_3 = d_3/\beta_1$ and $\nu_4 = d_4/\beta_2$.

Table 4.7

Class of Estimators for $t_{66}^{(i)}$

Class of Estimators	d_1	d_2	β_1	β_2
$t_{66}^{(1)} = \bar{z}\exp\left[\frac{\left(\bar{X}_1-\bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}\right)}{\bar{X}_1}\right]\exp\left[\frac{\left(\bar{X}_2-\bar{x}_2\right)}{\bar{X}_2}\right]$	1	1	1	1
$t_{66}^{(2)} = \bar{z}\exp\left[\frac{\left(\bar{X}_1-\bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}\right)}{2\left(\bar{X}_1+\bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}\right)}\right]\exp\left[\frac{\left(\bar{X}_2-\bar{x}_2\right)}{2\left(\bar{X}_2+\bar{x}_2\right)}\right]$	½	½	2	2
$t_{66}^{(3)} = \bar{z}\exp\left[\frac{\rho_{z_j x_1}\left(\bar{X}_1-\bar{x}_1\frac{\bar{X}_2}{\bar{x}_2}\right)}{\bar{X}_1}\right]\exp\left[\frac{\rho_{z_j x_2}\left(\bar{X}_2-\bar{x}_2\right)}{\bar{X}_2}\right]$	$\rho_{z_j x_1}$	$\rho_{z_j x_2}$	1	1

The bias and *MSE* of $t_{66}^{(i)}$, to the first order approximation $O\left(n^{-1}\right)$ may be given as,

$$Bias\left(t_{66}^{(i)}\right) \cong \bar{Z}_j \theta \left[\nu_3 \left(C_{z_j x_1} - C_{z_j x_2} \right) + \frac{1}{2} \nu_3^2 \left(C_{x_1}^2 + C_{x_2}^2 - 2C_{x_1 x_2} \right) - \nu_4 C_{z_j x_2} \right.$$

$$\left. - \nu_4 C_{z_j x_2} - \nu_3 \nu_4 \left(C_{x_1 x_2} - C_{x_2}^2 \right) + \frac{1}{2} \nu_4^2 C_{x_2}^2 \right], \quad (4.9.3)$$

and

$$MSE\left(t_{66}^{(i)}\right) \cong \theta \bar{Z}_j^2 \left[C_{z_j}^2 + \nu_3^2 \left(C_{x_1}^2 + C_{x_2}^2 - 2C_{x_1 x_2} \right) \right.$$

$$+ \nu_4^2 C_{x_2}^2 + 2\nu_3 \left(C_{z_j x_1} - C_{z_j x_2} \right)$$

$$\left. - 2\nu_4 C_{z_j x_2} - 2\nu_3 \nu_4 \left(C_{x_1 x_2} - C_{x_2}^2 \right) \right], \quad (4.9.4)$$

and, ν_3 and ν_4 attain their optimum values ν_3^{opt} and ν_4^{opt} respectively as,

$$\nu_3^{opt} = \frac{C_{z_j} K_4}{\left(K_5 + 3\left(C_{x_1} \rho_{x_1 x_2} - C_{x_2} \right) \right)},$$

and

$$\nu_4^{opt} = \frac{1}{C_{x_2}} \left[C_{z_j} \rho_{z_j x_2} - \nu_3 \left(C_{x_1} \rho_{x_1 x_2} - C_{x_2} \right) \right], \quad (4.9.5)$$

where

$$K_4 = 2\rho_{z_j x_2} \left(C_{x_1} \rho_{x_1 x_2} - C_{x_2} \right) - \left(C_{x_1} \rho_{z_j x_1} - C_{x_2} \rho_{z_j x_2} \right)$$

and

$$K_5 = C_{x_1}^2 + C_{x_2}^2 - 2C_{x_1 x_2}.$$

From (4.9.3- 4.9.4), we may get the bias and *MSE*'s for the estimator $t_{66}^{(i)}$ $(i = 1, 2, 3)$ using different choices of d_1, d_2, β_1 and β_2.Substituting the optimum values in (4.9.4), we may get the minimum value of the $MSE\left(t_{66}^{(i)}\right)$ as,

$$\min MSE\left(t_{66}^{(i)}\right) = \theta \bar{Z}_j^2 C_{z_j}^2 \left[1 - \frac{K_4^2}{\left(K_5 + 3\left(C_{x_1} \rho_{x_1 x_2} - C_{x_2} \right) \right)} \right]. \quad (4.9.6)$$

We may get min *MSE*s for $t_{66}^{(i)}$ $(i = 1,2,3)$ using different values of d_1, d_2, β_1 and β_2 in (4.9.6) (see as given in Table 4.7).

4.10 SIMULATION STUDY

In this section the efficiency of the proposed class of estimators over existing estimators using proposed model and Pollock and Bek (1976) is provided using real data and with a simulation study. This study is involved to evaluate the *MSE*s of the estimators both empirically and theoretically. This simulation study considers two finite populations of size *N*=1000 each generated from multivariate normal distribution with theoretical mean vector $\mu = [5,5,5]$ for $[Y, X_1, X_2]$ with covariance matrices for two populations respectively as,

Population I

$$\sigma^2 = \begin{bmatrix} 10 & 3 & 2.9 \\ 3 & 2 & 1.1 \\ 2.9 & 1.1 & 2 \end{bmatrix} \qquad \rho_{x1y} = 0.6817, \ \ \rho_{x2y} = 0.6705$$

Population II

$$\sigma^2 = \begin{bmatrix} 6 & 3 & 2.9 \\ 3 & 2 & 1.1 \\ 2.9 & 1.1 & 2 \end{bmatrix} \qquad \rho_{x1y} = 0.8706, \ \ \rho_{x2y} = 0.8428$$

The scrambling variables *S* and *R* are taken from the normal distribution with mean zero and standard deviation equal to 10% of the standard deviation of X_1.

The sample size for each population considered as $n = 50{,}100$ and 300.

In Table (4.8-4.12), the results of *MSE*'s of the estimators for both populations have been presented. The empirical and theoretical results are obtained for additive model and proposed *SRR* models using $a = 0.5$ and $g = 0.6$.

4.11 DISCUSSION

In this chapter, we have proposed some generalized estimators to estimate population mean of the sensitive variable y under the observed response z_j, using two auxiliary variables. The expressions of the bias and *MSE* have been derived. The performance of the estimators have been examined using two data sets and compare the results of proposed *SRR* models and the additive model presented by Pollock and Bek (1971). From Table (4.8-4.13), it is observed as the sample size increases from 50 to 300 we obtained minimum *MSE*'s of the estimators. Also, the proposed estimators perform better than existing estimator t_{45}. Furthermore, the generalized estimator $t_{62}^{(i)}$ is more efficient among all the estimators presented in the Table (4.8-4.13). It is observed that the *MSE*'s obtained are minimum using proposed *SRR* models as compare to the *MSE*'s obtained using additive model.

Table 4.8
Simulation results at sample size n=50 for the MSEs of the Estimators using Population I [N=1000, $\rho_{x1y} = 0.6817$ & $\rho_{x2y} = 0.6705$]

n=50		**MSE Estimation using Population–I**									
		Pollock and Bek Model		**Proposed Models $g = 0.6$ & $a = 0.5$**							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Estimators		**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**
t_m		0.61963	0.61033	0.50974	0.51082	0.52077	0.51049	0.52034	0.51071	0.52036	0.51024
t_{K_2}		0.89517	0.81031	0.71880	0.71146	0.79966	0.75127	0.74865	0.75085	0.79800	0.78111
Proposed Estimators	$t_{59}^{(i)}$	0.35519	0.34310	0.30526	0.30591	0.31520	0.31535	0.31576	0.31557	0.34881	0.34522
	$t_{60}^{(i)}$	0.35338	0.35606	0.29583	0.30578	0.31510	0.30524	0.30556	0.31546	0.33204	0.33512
	$t_{61}^{(i)}$	0.35184	0.34310	0.29559	0.30591	0.31498	0.31535	0.30159	0.31557	0.32370	0.34522
	$t_{62}^{(i)}$	0.31134	0.35606	0.29539	0.29578	0.31481	0.30524	0.30501	0.31546	0.31490	0.33512
	t_{63}	0.37650	0.39361	0.31212	0.31171	0.32018	0.32642	0.31196	0.32118	0.32421	0.32114
	$t_{64}^{(i)}$	0.32101	0.31105	0.30853	0.30240	0.31320	0.31206	0.30750	0.31051	0.30339	0.30314
	$t_{65}^{(i)}$	0.31651	0.30361	0.30397	0.30432	0.30907	0.31409	0.30425	0.31042	0.30354	0.30419
	$t_{66}^{(i)}$	0.31011	0.30196	0.30049	0.30055	0.30856	0.31199	0.30103	0.31020	0.30119	0.30194

Table 4.9
Simulation results at sample size n=100 for the MSEs of the Estimators using Population I [N=1000, $\rho_{x1y}=0.6817$ & $\rho_{x2y}=0.6705$]

n=100		MSE Estimation using Population–I									
		Pollock and Bek Model		Proposed Models $g=0.6$ & $a=0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Estimators		**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**
t_m		0.28315	0.28956	0.21015	0.21126	0.26104	0.26711	0.25090	0.25100	0.27942	0.27113
t_{K_2}		0.42122	0.41089	0.31968	0.31443	0.38202	0.35761	0.33755	0.34073	0.40331	0.40110
Proposed Estimators	$t_{59}^{(i)}$	0.21912	0.21372	0.15662	0.16587	0.18588	0.18256	0.18496	0.18164	0.20552	0.20782
	$t_{60}^{(i)}$	0.21597	0.21512	0.15237	0.16315	0.18541	0.17542	0.17993	0.17285	0.20508	0.20752
	$t_{61}^{(i)}$	0.21404	0.21372	0.14816	0.16587	0.17354	0.18256	0.17284	0.18164	0.20475	0.20782
	$t_{62}^{(i)}$	0.21098	0.21512	0.15010	0.16315	0.16992	0.17542	0.16807	0.17285	0.20433	0.20752
	t_{63}	0.23250	0.24974	0.20592	0.20567	0.22504	0.21896	0.20631	0.20106	0.22056	0.22199
	$t_{64}^{(i)}$	0.22728	0.23996	0.20182	0.20775	0.21178	0.20985	0.20342	0.19971	0.21517	0.21015
	$t_{65}^{(i)}$	0.22172	0.21697	0.20520	0.20187	0.20147	0.20196	0.20190	0.20089	0.21198	0.21226
	$t_{66}^{(i)}$	0.22045	0.21860	0.20140	0.20964	0.20444	0.20972	0.20057	0.20014	0.21053	0.21020

Table 4.10
Simulation results at sample size n=300 for the MSEs of the Estimators using Population I [N=1000, $\rho_{x1y} = 0.6817$ & $\rho_{x2y} = 0.6705$]

n=300		**MSE Estimation using Population–I**									
		Pollock and Bek Model		**Proposed Models** $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Estimators		**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**
t_m		0.03874	0.03281	0.02580	0.02511	0.03405	0.034101	0.02930	0.02962	0.03524	0.03519
t_{K_2}		0.09553	0.09113	0.07996	0.07869	0.08950	0.08784	0.08588	0.84786	0.08959	0.08910
Proposed Estimators	$t_{59}^{(i)}$	0.02638	0.02688	0.02417	0.02493	0.02521	0.02544	0.02553	0.02501	0.25628	0.25105
	$t_{60}^{(i)}$	0.02609	0.02687	0.02388	0.02398	0.02519	0.02543	0.02610	0.02577	0.25572	0.25835
	$t_{61}^{(i)}$	0.02538	0.02688	0.02408	0.02493	0.02518	0.02544	0.02467	0.02501	0.2525	0.25105
	$t_{62}^{(i)}$	0.02509	0.02687	0.02399	0.02398	0.02516	0.02543	0.02455	0.02577	0.25025	0.25835
	t_{63}	0.03073	0.03157	0.02939	0.02972	0.02952	0.02852	0.03054	0.30015	0.03048	0.03015
	$t_{64}^{(i)}$	0.03040	0.03050	0.02828	0.02819	0.02810	0.02718	0.02954	0.02952	0.02961	0.02926
	$t_{65}^{(i)}$	0.29081	0.02995	0.02848	0.02694	0.02816	0.02747	0.02952	0.28228	0.02905	0.20497
	$t_{66}^{(i)}$	0.29097	0.02974	0.02812	0.02429	0.02875	0.02758	0.02914	0.28791	0.02546	0.02012

Table 4.11
Simulation results at sample size n=50 for the MSEs of the Estimators using Population II [N=1000, $\rho_{x1y} = 0.8706$ & $\rho_{x2y} = 0.8428$]

n=50		MSE Estimation using Population–II									
		Pollock and Bek Model		**Proposed Models** $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Estimators		**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**
t_m		0.11751	0.11601	0.09403	0.09600	0.11011	0.11003	0.10125	0.10655	0.10134	0.10681
t_{K_2}		0.80170	0.81284	0.60321	0.60121	0.63395	0.64481	0.75579	0.71065	0.68878	0.61071
Proposed Estimators	$t_{59}^{(i)}$	0.10473	0.10462	0.04684	0.04529	0.06310	0.06511	0.04767	0.04737	0.09535	0.09487
	$t_{60}^{(i)}$	0.10461	0.10454	0.04558	0.04448	0.06241	0.06392	0.04637	0.04653	0.09520	0.09479
	$t_{61}^{(i)}$	0.10440	0.10462	0.04277	0.04529	0.06134	0.06511	0.04428	0.04737	0.09483	0.09487
	$t_{62}^{(i)}$	0.10428	0.10454	0.04165	0.04448	0.06030	0.06392	0.04310	0.04653	0.09470	0.09479
	t_{63}	0.11034	0.11032	0.05340	0.05384	0.06936	0.06515	0.05476	0.05409	0.10046	0.09452
	$t_{64}^{(i)}$	0.10850	0.10618	0.05177	0.05075	0.06744	0.06208	0.05405	0.05318	0.09791	0.09148
	$t_{65}^{(i)}$	0.10781	0.10813	0.05142	0.05031	0.06655	0.06115	0.05436	0.05350	0.09802	0.99336
	$t_{66}^{(i)}$	0.10720	0.10814	0.05122	0.05011	0.06615	0.06110	0.05418	0.05312	0.09350	0.09872

Table 4.12
Simulation results at sample size n=100 for the MSEs of the Estimators using Population II [N=1000, $\rho_{x1y} = 0.8706$ & $\rho_{x2y} = 0.8428$]

n=100		MSE Estimation using Population–II									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Estimators		**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**
t_m		0.05818	0.05942	0.04567	0.04633	0.05293	0.05360	0.05196	0.51250	0.05159	0.05283
t_{K_2}		0.34975	0.34121	0.22486	0.22183	0.34261	0.33204	0.30391	0.30116	0.29681	0.29117
Proposed Estimators	$t_{59}^{(i)}$	0.02261	0.02167	0.01830	0.01872	0.02130	0.02131	0.02068	0.02018	0.02189	0.02144
	$t_{60}^{(i)}$	0.02232	0.02148	0.01523	0.01574	0.02101	0.02109	0.02041	0.02012	0.02187	0.02126
	$t_{61}^{(i)}$	0.02190	0.02167	0.01774	0.01872	0.02106	0.02131	0.02001	0.02018	0.02182	0.02144
	$t_{62}^{(i)}$	0.02162	0.02148	0.01731	0.01574	0.02015	0.02109	0.01995	0.02012	0.02179	0.02126
	t_{63}	0.02276	0.02215	0.01912	0.01916	0.02190	0.02225	0.02187	0.2201	0.02219	0.02214
	$t_{64}^{(i)}$	0.02250	0.02187	0.01916	0.01905	0.02124	0.02210	0.02153	0.02159	0.02212	0.02177
	$t_{65}^{(i)}$	0.02155	0.02196	0.01910	0.01897	0.02133	0.02130	0.02124	0.02153	0.02215	0.02165
	$t_{66}^{(i)}$	0.02193	0.02200	0.01895	0.01890	0.02124	0.02161	0.02132	0.02154	0.02199	0.02150

Table 4.13
Simulation results at sample size n=300 for the MSEs of the Estimators using Population II [N=1000, $\rho_{x1y}=0.8706$ & $\rho_{x2y}=0.8428$]

n=300		MSE Estimation using Population–II									
		Pollock and Bek Model		Proposed Models $g=0.6$ & $a=0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
t_m		0.01449	0.01593	0.01377	0.13590	0.01444	0.01455	0.01401	0.01406	0.01415	0.01465
t_{K_2}		0.10094	0.11309	0.07357	0.07105	0.09512	0.09567	0.07593	0.07432	0.08942	0.08811
Proposed Estimators	$t_{59}^{(i)}$	0.00555	0.00560	0.00415	0.00406	0.00507	0.00575	0.00535	0.00546	0.00552	0.00556
	$t_{60}^{(i)}$	0.00553	0.00559	0.00413	0.00415	0.00505	0.00574	0.00533	0.00503	0.00548	0.00543
	$t_{61}^{(i)}$	0.00550	0.00560	0.00408	0.00406	0.00502	0.00575	0.00529	0.00546	0.00535	0.00506
	$t_{62}^{(i)}$	0.00548	0.00559	0.00406	0.00415	0.00495	0.00574	0.00527	0.00503	0.00511	0.00543
	t_{63}	0.01039	0.01037	0.00496	0.00472	0.00806	0.00818	0.00946	0.00953	0.00952	0.00904
	$t_{64}^{(i)}$	0.00729	0.00711	0.00447	0.00477	0.00637	0.00639	0.00546	0.00555	0.00623	0.00691
	$t_{65}^{(i)}$	0.00766	0.00701	0.00437	0.00459	0.00625	0.00618	0.00538	0.00542	0.00636	0.00640
	$t_{66}^{(i)}$	0.00784	0.00778	0.00426	0.00437	0.00662	0.00615	0.00525	0.00514	0.00634	0.00615

CHAPTER 5

PROPOSED ESTIMATORS TO ESTIMATE POPULATION MEAN FOR TWO-PHASE SAMPLING BASED ON SCRAMBLED RANDOMIZED RESPONSE MODELS

5.1 INTRODUCTION

This chapter presents some estimators under two-phase sampling using non-sensitive auxiliary variables related to sensitive study variable *y*. The coding mechanism z_j (j=4, G1, G2, G3, G4) is used to estimate population mean of *y*. The bias and the *MSE*'s of the proposed estimators has been derived up to first order of approximation. The estimators are presented in three different cases of two-phase sampling introduced by Sammiuddin and Hanif (2007), regarding the availability of auxiliary information on more than one auxiliary variables such as i.e., full-information case, partial information case or no-information case.

5.2 ESTIMATORS PROPOSED FOR FULL-INFORMATION CASE (FIC)

In this section using full-information case (*FIC*), some estimators have been presented for two-phase sampling based on scrambled models. To estimate these estimators, it is considered that the population mean of both non-sensitive auxiliary variables say x_1 and x_2 are available in prior. Also, assume that the auxiliary information on non-sensitive variable x_1 is collected at the first-phase and another non-sensitive auxiliary variable x_2 and the observed response *z* are collected at second-phase respectively.

Some generalized estimators are also suggested based on scrambled responses. Further, the bias and mean square error of the estimators have been derived and the class of estimators are also given.

5.2.1 Proposed Estimator-I

Following Gupta et al. (2012), regression estimator using two non-sensitive auxiliary variables have been presented for population mean under two-phase sampling. The estimator is as,

$$t_{67} = \bar{z}''_j + b_1\left(\bar{X}_1 - \bar{x}''_1\right) + b_2\left(\bar{X}_2 - \bar{x}''_2\right), \text{ for } j = 4, \text{G1, G2, G3, G4} \quad (5.2.1.1)$$

where b_1 and b_2 are the sample regression coefficients between z_j and x_1, and z_j and x_2 respectively, scrambled on *y*. To the first order of approximation, the estimator t_{67} is unbiased, so to find the variance (5.2.1.1) may be expanded in terms of *e*'s, estimator t_{67} becomes,

$$t_{67} = \bar{Z}_j\left(1+e''_{\bar{z}_j}\right) - b_1\bar{X}_1 e''_{\bar{x}_1} - b_2\bar{X}_2 e''_{\bar{x}_2}. \tag{5.2.1.2}$$

Squaring on both sides we may get

$$\left(t_{67}-\bar{Z}\right)^2 = \bar{Z}_j^2 e''^2_{\bar{z}_j} + b_1^2\bar{X}_1^2 e''^2_{\bar{x}_1} + b_1^2\bar{X}_2^2 e''^2_{\bar{x}_2} - 2\bar{Z}_j\bar{X}_1 b_1 e''_{\bar{z}_j} e''_{\bar{x}_1}$$
$$-2\bar{Z}_j\bar{X}_2 b_2 e''_{\bar{z}_j} e''_{\bar{x}_2} + 2\bar{X}_1\bar{X}_2 b_1 b_2 e''_{\bar{x}_1} e''_{\bar{x}_2}. \tag{5.2.1.3}$$

Applying expectation both sides, we may obtain,

$$MSE\left(t_{67}\right) = E_D\left(t_{67}-\bar{Z}_j\right)^2 = \theta_2\Big[\bar{Z}_j{}^2 C_Z^2 + b_1^2\bar{X}_1^2 C_{X1}^2 + b_1^2\bar{X}_2^2 C_{X2}^2$$
$$-2\bar{Z}_j\bar{X}_1 b_1 C_{ZX1} - 2\bar{Z}_j\bar{X}_2 b_2 C_{ZX2} + 2\bar{X}_1\bar{X}_2 b_1 b_2 C_{X1X2}\Big]. \tag{5.2.1.4}$$

The $MSE\left(t_{67}\right)$ is minimum for the optimum values of sample regression coefficients b_1 and b_2 as,

$$b_1^{opt} = \frac{\bar{Z}_j}{\bar{X}_1}\frac{C_{z_j}\left(\rho_{z_j x_1} - \rho_{z_j x_2}\rho_{x\,x_2}\right)}{C_{x_1}\left(1-\rho^2_{x_1x_2}\right)} = \frac{\bar{Z}_j}{\bar{X}_1}\frac{C_{z_j}}{C_{x_1}}\frac{A_1}{A_2}(Say), \tag{5.2.1.5}$$

and

$$b_2^{opt} = \frac{\bar{Z}_j}{\bar{X}_2}\frac{C_{z_j}\left(\rho_{Zx_2} - \rho_{Zx_1}\rho_{x_1x_2}\right)}{C_{x_2}\left(1-\rho^2_{x_1x_2}\right)} = \frac{\bar{Z}_j}{\bar{X}_2}\frac{C_{z_j}}{C_{x_2}}\frac{A_3}{A_2}(Say). \tag{5.2.1.6}$$

The minimum value of $MSE\left(t_{67}\right)$ using optimum values (5.2.1.5-5.2.1.6) may be obtained as,

$$MSE\left(t_{67}\right) = \bar{Z}_j^2\theta_2 C_{z_j}^2\Bigg[1 + \frac{1}{A_2^3}\left(A_1^2 + A_3^2 + 2\rho_{x_1x_2}A_1A_3\right)$$
$$-2\frac{1}{A_2}\left(A_1\rho_{z_jx_1} + A_3\rho_{z_jx_2}\right)\Bigg]. \tag{5.2.1.7}$$

5.2.2 Proposed Estimator-II

Motivated by Sousa et al. (2010), a ratio estimator under two-phase sampling based on scrambled response *y* have been proposed when observed response have positive correlation with non-sensitive variables x_1 and x_2. The proposed ratio estimator t_{68} is given as,

$$t_{68} = \bar{z}''_j \frac{\bar{X}_1}{\bar{x}''_1} \frac{\bar{X}_2}{\bar{x}''_2}, \qquad \text{for } j = 4, \text{G1, G2, G3, G4.} \tag{5.2.2.1}$$

To obtain the bias and mean square error of t_{68}, (5.2.2.1) may be expressed as,

$$t_{68} = \bar{Z}_j \left(1 + e''_{\bar{z}_j}\right)\left(1 + e''_{\bar{x}_1}\right)^{-1}\left(1 + e''_{\bar{x}_2}\right)^{-1}, \tag{5.2.2.2}$$

or

$$t_{68} - \bar{Z}_j = \bar{Z}_j \left[e''_{\bar{z}_j} - e''_{\bar{x}_1} + e''^2_{\bar{x}_1} - e''_{\bar{z}_j} e''_{\bar{x}_1} - e''_{\bar{x}_2} - e''_{\bar{z}_j} e''_{\bar{x}_2} + e''_{\bar{x}_1} e''_{\bar{x}_2} + e''^2_{\bar{x}_1} \right]. \tag{5.2.2.3}$$

Applying expectation on both sides, we may get the bias of t_{68} as,

$$Bias\left(t_{68}\right) = \bar{Z}_j \theta_2 \left[C^2_{x_1} + C^2_{x_2} - C_{z_j x_1} - C_{z_j x_2} + C_{x_1 x_2} \right]. \tag{5.2.2.4}$$

Taking (5.2.10) up to first order of approximation and squaring on both sides we may obtain,

$$\left(t_{68} - \bar{Z}_j\right)^2 = \bar{Z}^2_j \left[e''^2_{z_j x_1} + e''^2_{x_1} + e''^2_{x_2} - 2e''_{x_1} e''_{z_j} - 2e''_{x_2} e''_{z_j} + 2e''_{x_1} e''_{x_2} \right]. \tag{5.2.2.5}$$

Now taking expectation both sides we may have the $MSE\left(t_{68}\right)$ as,

$$MSE\left(t_{68}\right) = \bar{Z}^2_j \theta_2 \left[C^2_{z_j} + C^2_{x_1} + C^2_{x_2} - 2C_{z_j x_1} - 2C_{z_j x_2} + 2C_{x_1 x_2} \right]. \tag{5.2.2.6}$$

5.2.3 Proposed Estimator-III

Motivated from Singh and Vishwakarma (2007), ratio-type exponential estimator under two-phase sampling have been proposed using two auxiliary variables based on observed response z. The estimator is as,

$$t_{69} = \bar{z}''_j \exp\left[\frac{\bar{X}_1 - \bar{x}''_1}{\bar{X}_1 + \bar{x}''_1}\right] \exp\left[\frac{\bar{X}_2 - \bar{x}''_2}{\bar{X}_2 + \bar{x}''_2}\right],$$

for *j*= 4, G1, G2, G3, G4. (5.2.3.1)

Rewrite (5.2.3.1) in the form of *e*'s as,

$$t_{69} = \bar{Z}_j\left(1 + e''_{\bar{z}_j}\right)\exp\left[\frac{-e''_{\bar{x}_1}}{2}\left(1 + \frac{e''_{\bar{x}_1}}{2}\right)^{-1}\right]\exp\left[\frac{-e''_{\bar{x}_2}}{2}\left(1 + \frac{e''_{\bar{x}_2}}{2}\right)^{-1}\right], \quad (5.2.3.2)$$

and solving (5.2.3.2), we may have,

$$t_{69} - \bar{Z}_j = \bar{Z}_j\left[e''_{\bar{z}_j} - \frac{e''_{\bar{x}_1}}{2} + \frac{3e''^2_{\bar{x}_1}}{8} - \frac{e''_{\bar{x}_2}}{2} + \frac{e''_{\bar{x}_1}e''_{\bar{x}_2}}{4} + \frac{3e''^2_{\bar{x}_2}}{8} - \frac{e''_{\bar{x}_1}e''_{\bar{z}_j}}{2} - \frac{e''_{\bar{x}_2}e''_{\bar{z}_j}}{2}\right].$$

(5.2.3.3)

The expressions of the bias and *MSE* of t_{69} are obtained as,

$$Bias\left(t_{69}\right) = \frac{1}{2}\bar{Z}_j\theta_2\left[\frac{3}{8}\left(C^2_{x_1} + C^2_{x_2}\right) - C_{z_jx_1} + \frac{1}{2}C_{x_1x_2} - C_{z_jx_2}\right], \quad (5.2.3.4)$$

and

$$MSE\left(t_{69}\right) = \bar{Z}^2_j\theta_2\left[C^2_{z_j} + \frac{1}{4}\left(C^2_{x_1} + C^2_{x_2}\right) + \frac{1}{2}C_{x_1x_2} - \left(C_{z_jx_1} + C_{z_jx_2}\right)\right]. \quad (5.2.3.5)$$

5.2.4 Proposed Estimator-IV

The proposed estimator has been generalized by following Koyuncu et al. (2014) regression-type exponential estimator under two-phases sampling. The generalized exponential-type estimator is given by,

$$t^{(i)}_{70} = \left[\alpha_0\bar{z}''_j + \alpha_1\left(\bar{X}_2 - \bar{x}''_2\right)\right]\exp\left[\frac{a_1\left(\bar{X}_1 - \bar{x}''_1\right)}{\bar{X}_1 + \left(\alpha_2 - 1\right)\bar{x}''_1}\right],$$

for *j*= 4, G1, G2, G3, G4 (5.2.4.1)

Table 5.1
Class of Estimators for $t_{70}^{(i)}$

Class of Estimators	α_0	α_1	a_1	α_2
$t_{70}^{(1)} = \left(\alpha_0 \bar{z}_j'' + \left(\bar{X}_2 - \bar{x}_2''\right)\right)\left[\exp\left(\frac{\left(\bar{X}_1 - \bar{x}_1''\right)}{\bar{X}_1}\right)\right]$	α_0	1	1	1
$t_{70}^{(2)} = \left(\alpha_0 \bar{z}_j'' + b_{z_j x_2}\left(\bar{X}_2 - \bar{x}_2''\right)\right)\left[\exp\left(\frac{\left(\bar{X}_1 - \bar{x}_1''\right)}{\bar{X}_1}\right)\right]$	α_0	$b_{z_j x_2}$	1	1
$t_{70}^{(3)} = \left(\alpha_0 \bar{z}_j'' + \rho_{z_j x_2}\left(\bar{X}_2 - \bar{x}_2''\right)\right)\left[\exp\left(\frac{\left(\bar{X}_1 - \bar{x}_1''\right)}{\bar{X}_1}\right)\right]$	α_0	$\rho_{z_j x_2}$	1	1
$t_{70}^{(4)} = \left(\alpha_0 \bar{z}_j'' + \left(\bar{X}_2 - \bar{x}_2''\right)\right)\left[\exp\left(\frac{\bar{X}_1 - \bar{x}_1''}{2\left(\bar{X}_1 + \bar{x}_1''\right)}\right)\right]$	α_0	1	½	2
$t_{70}^{(5)} = \left(\alpha_0 \bar{z}_j'' + b_{z_j x_2}\left(\bar{X}_2 - \bar{x}_2''\right)\right)\left[\exp\left(\frac{\bar{X}_1 - \bar{x}_1''}{2\bar{X}_1}\right)\right]$	α_0	$b_{z_j x_2}$	½	1
$t_{70}^{(6)} = \left(\alpha_0 \bar{z}_j'' + C_{x_2}\left(\bar{X}_2 - \bar{x}_2''\right)\right)\left[\exp\left(\frac{2\left(\bar{X}_1 - \bar{x}_1''\right)}{\bar{X}_1}\right)\right]$	α_0	C_{x2}	2	1
$t_{70}^{(7)} = \left(\alpha_0 \bar{z}_j'' + \rho_{z_j x_2}\left(\bar{X}_2 - \bar{x}_2''\right)\right)\left[\exp\left(\frac{\rho_{z_j x_1}\left(\bar{X}_1 - \bar{x}_1''\right)}{\bar{X}_1}\right)\right]$	α_0	$\rho_{z_j x_2}$	$\rho_{z_j x_1}$	1
$t_{70}^{(8)} = \left(\bar{z}_j'' + b_{z_j x_2}\left(\bar{X}_2 - \bar{x}_2''\right)\right)\left[\exp\left(\frac{\left(\bar{X}_1 - \bar{x}_1''\right)}{\bar{X}_1}\right)\right]$	1	$b_{z_j x_2}$	1	1
$t_{70}^{(9)} = \left(\bar{z}_j'' + \rho_{z_j x_2}\left(\bar{X}_2 - \bar{x}_2''\right)\right)\left[\exp\left(\frac{\left(\bar{X}_1 - \bar{x}_1''\right)}{\bar{X}_1}\right)\right]$	1	$\rho_{z_j x_2}$	1	1

where $a_1(0,-1,+1)$ is a generalization constant whose values yield regression estimator, class of regression-cum-exponential ratio and regression-cum-exponential product estimators respectively. $\alpha_{0,}$ α_1, and α_2 are assumed to be unknown constants and need to be estimated such that the *MSE* of $t_{70}^{(i)}$ is minimum. For different choices of $\alpha_{0,}$ α_1, α_2 and a_1, we may get different estimators presented below in Table 5.1.

To obtain the bias and the *MSE* of $t_{70}^{(i)}$, (5.2.4.1) may be expressed in the form of e's as,

$$t_{70}^{(i)}=\left[\alpha_0\bar{Z}_j\left(1+e''_{\bar{z}_j}\right)-\alpha_1\bar{X}_2e''_{\bar{x}_2}\right]\exp\left[\frac{-a_1e''_{\bar{x}_1}}{\alpha_2\left(1+e''_{\bar{x}_1}\right)}\left(1-\frac{e''_{\bar{x}_1}}{\alpha_2\left(1+e''_{\bar{x}_1}\right)}\right)^{-1}\right], \tag{5.2.4.2}$$

and solving (5.2.4.2), we may have,

$$t_{70}^{(i)}-\bar{Z}_j=\bar{Z}_j\left(\alpha_0-1\right)+\alpha_0\bar{Z}_j\left[e''_{\bar{z}_j}-w_1e''_{\bar{x}_1}+\frac{1}{2}w_1^2e''^{\,2}_{\bar{x}_1}-w_1e''_{\bar{z}_j}e''_{\bar{x}_1}\right]$$
$$-\alpha_1\bar{X}_2\left[e''_{\bar{x}_2}-w_1e''_{\bar{x}_2}e''_{\bar{x}_1}\right], \tag{5.2.4.3}$$

where $w_1=a_1/\alpha_2$. The expressions of the bias and *MSE* are obtained as,

$$Bias\left(t_{70}^{(i)}\right)=\bar{Z}_j\left(\alpha_0-1\right)+\alpha_0\bar{Z}_j\theta_2w_1C_{x_1}\left[\frac{1}{2}w_1C_{x_1}-\theta_3w_1C_{z_j}\rho_{z_jx_1}\right]$$
$$+\alpha_1\bar{X}_2\theta_2w_1C_{x_1x_2}, \tag{5.2.4.4}$$

and

$$MSE\left(t_{70}^{(i)}\right)=\bar{Z}_j^2\left(\alpha_0-1\right)^2+\alpha_0^2\bar{Z}_j^2\theta_2\left[C_{z_j}^2+w_1^2C_{x_1}^2-2w_1C_{z_jx_1}\right]$$
$$+\alpha_1^2\bar{X}_2^2\theta_2C_{x_2}^2+2\bar{Z}_j^2\left(\alpha_0-1\right)\alpha_0\theta_2\left[\frac{1}{2}w_1^2C_{x_1}^2-w_1C_{z_jx_1}\right]$$
$$-2\bar{Z}_j\bar{X}_2\left(\alpha_0-1\right)\alpha_1w_1\theta_2C_{x_1x_2}+2\alpha_0\alpha_1\bar{Z}_j\bar{X}_2\theta_2\left[C_{z_jx_2}-w_1C_{x_1x_2}\right]. \tag{5.2.4.5}$$

The optimum values of w_1,α_0 and α_1 are obtained as,

$$w_1^{opt}=\frac{\left(\bar{Z}_jC_{z_j}\rho_{z_jx_1}-\alpha_1C_{x_1}\rho_{x_1x_2}\right)}{\bar{Z}_jC_{z_j}},\ \alpha_1^{opt}=\frac{\alpha_0\bar{Z}_jC_{z_j}\left(\rho_{z_jx_2}-\rho_{z_jx_1}\rho_{x_1x_2}\right)}{\bar{X}_2C_{x_2}\left(1-\rho_{x_1x_2}^2\right)}, \tag{5.2.4.6}$$

and

$$\alpha_0^{opt}=\frac{1}{1+H_1C_{z_j}^2}, \tag{5.2.4.7}$$

where

$$H_1 = \left[\begin{array}{l} 1+\dfrac{\left(\rho_{z_j x_2}-\rho_{z_j x_1}\rho_{x_1 x_2}\right)^2}{\left(1-\rho_{x_1 x_2}^2\right)^2}+\dfrac{\left(\rho_{z_j x_1}-\rho_{z_j x_2}\rho_{x_1 x_2}\right)^2}{\left(1-\rho_{x_1 x_2}^2\right)^2} \\ \qquad\qquad -2\dfrac{\left(\rho_{z_j x_2}-\rho_{z_j x_1}\rho_{x_1 x_2}\right)}{\left(1-\rho_{x_1 x_2}^2\right)}-\dfrac{\left(\rho_{z_j x_2}-\rho_{z_j x_1}\rho_{x_1 x_2}\right)}{\left(1-\rho_{x_1 x_2}^2\right)} \end{array}\right].$$

We may get the bias and the MSE for $t_{70}^{(i)}$ (i=1,2,…,9) using different values of $\alpha_{0,}\,\alpha_1$, α_2 and a_1 in (5.2.4.4-5.2.4.5) (e.g. as given in Table 5.1). Now, the minimum MSE of $t_{70}^{(i)}$ may be obtained as,

$$\min MSE\left(t_{70}^{(i)}\right)=\bar{Z}_j^2\theta_2 C_{z_j}^2\left[1-\frac{1}{H_1}\right]. \qquad (5.2.4.8)$$

We may get the min MSE for $t_{70}^{(i)}$ (i=1,2,…,9) using different values of $\alpha_{0,}\,\alpha_1$, α_2 and a_1 in (5.2.4.8) as given in Table 5.1.

5.3 ESTIMATORS PROPOSED FOR PARTIAL-INFORMATION CASE (PIC)

In this section, some regression, ratio and exponential-type estimators based on SRR models to estimate population mean of sensitive study variable y have been advised under partial information case of two-phase sampling. It is assumed that the information on non-sensitive auxiliary variable x_2 is already known whereas the information on non-sensitive auxiliary variable x_1 is not available in prior, in such situation, the sample mean $\bar{x}_1'$ is used to estimate from first-phase sample of size n_1. The information at second-phase are collected for non-sensitive auxiliary variable x_2 and the sensitive variable under study say y are collected at second-phase respectively. The bias and mean squared error for proposed estimators have been derived.

5.3.1 Proposed Estimator-V

The regression estimator for partial-information case using two non-sensitive auxiliary variables under scrambles response y is presented as,

$$t_{71} = \bar{z}''_j + b_3\left(\bar{x}_1{}' - \bar{x}''_1\right) + b_4\left(\bar{X}_2 - \bar{x}''_2\right),$$

for j= 4, G1, G2, G3, G4. (5.3.1.1)

where b_3 and b_4 are the sample regression coefficients between z_{j} and x_1, and z_{j} and x_2 respectively scrambled on y.

Expanding (5.3.1.1) in terms of e's, we may have t_{71} as,

$$t_{71} = \bar{Z}_j\left(1 + e''_{\bar{z}_j}\right) + b_3\bar{X}_1\left(e'_{\bar{x}1} - e''_{\bar{x}1}\right) - b_4\bar{X}_2 e''_{\bar{x}2}. \quad (5.3.1.2)$$

Squaring on both sides we may get,

$$\begin{aligned}\left(t_{71} - \bar{Z}_j\right)^2 &= \bar{Z}_j^2 e''^2_{\bar{z}_j} + b_3^2\bar{X}_1^2\left(e'^2_{\bar{x}1} + e''^2_{\bar{x}1} - 2e'_{\bar{x}1}e''_{\bar{x}1}\right) \\ &+ b_4^2\bar{X}_2^2 e''^2_{\bar{x}2} + 2\bar{Z}_j\bar{X}_1 b_3\left(e'_{\bar{x}1}e''_{\bar{z}_j} - e''_{\bar{x}1}e''_{\bar{z}_j}\right) \\ &- 2\bar{Z}_j\bar{X}_2 b_4 e''_{\bar{z}_j}e''_{\bar{x}2} - 2\bar{Z}_j\bar{X}_1\bar{X}_2 b_3 b_4\left(e'_{\bar{x}1}e''_{\bar{x}2} - e''_{\bar{x}1}e''_{\bar{x}2}\right), \quad (5.3.1.3)\end{aligned}$$

Applying expectation both sides, we may obtain,

$$\begin{aligned}MSE\left(t_{71}\right) &= \bar{Z}_j^2\theta_2 C_{z_j}^2 + b_3^2\bar{X}_1^2\left(\theta_1 C_{x_1}^2 + \theta_2 C_{x_1}^2 - 2\theta_1 C_{x_1}^2\right) \\ &+ b_4^2\bar{X}_2^2\theta_2 C_{x_2}^2 + 2\bar{Z}_j\bar{X}_1 b_3\left(\theta_1 C_{z_j x_1} - \theta_2 C_{z_j x_1}\right) \\ &- 2\bar{Z}_j\bar{X}_2 b_4\theta_2 C_{z_j x_2} - 2\bar{Z}_j\bar{X}_1\bar{X}_2 b_3 b_4\left(\theta_1 C_{x_1 x_2} - \theta_2 C_{x_1 x_2}\right),\end{aligned}$$

or

$$\begin{aligned}MSE\left(t_{71}\right) &= \bar{Z}_j^2\theta_2 C_{z_j}^2 + b_3^2\bar{X}_1^2\theta_3 C_{x_1}^2 + b_4^2\bar{X}_2^2\theta_2 C_{x_2}^2 - 2\bar{Z}_j\bar{X}_1 b_3\theta_3 C_{z_j x_1} \\ &- 2\bar{Z}_j\bar{X}_2 b_4\theta_2 C_{z_j x_2} + 2\bar{Z}_j\bar{X}_1\bar{X}_2 b_3 b_4\theta_3 C_{x_1 x_2}. \quad (5.3.1.4)\end{aligned}$$

The $MSE\left(t_{71}\right)$ is minimum for the optimum values of the regression coefficients b_1 and b_2 are as,

$$b_3^{opt} = \frac{\bar{Z}_j}{\bar{X}_1} \frac{\theta_2 C_{z_j}\left(\rho_{z_j x_1} - \rho_{z_j x_2}\rho_{x_1 x_2}\right)}{C_{x_1}\left(\theta_2 - \theta_3 \rho_{x_1 x_2}^2\right)} = \frac{\bar{Z}_j}{\bar{X}_1} \frac{\theta_2 C_{z_j}}{C_{x_1}} \frac{A_4}{A_5} (Say), \qquad (5.3.1.5)$$

and

$$b_4^{opt} = \frac{\bar{Z}_j}{\bar{X}_1} \frac{C_{z_j}\left(\theta_2 \rho_{z_j x_2} - \theta_2 \rho_{z_j x_1}\rho_{x_1 x_2}\right)}{C_{X2}\left(\theta_2 - \theta_3 \rho_{x_1 x_2}^2\right)} = \frac{\bar{Z}_j}{\bar{X}_1} \frac{\theta_2 C_{xz}}{C_{x_2}} \frac{A_6}{A_5} (Say), \qquad (5.3.1.6)$$

and the minimum value of $MSE\left(t_{71}\right)$ using optimum values (5.3.1.5-5.3.1.6) is

$$MSE\left(t_{71}\right) = \bar{Z}_j^2 C_{z_j}^2 \left[\begin{array}{c} \theta_2\left(1 + \frac{A_6^2}{A_5^3} - 2\frac{A_6}{A_5}\rho_{z_j x_2}\right) \\ + \theta_3\left(\frac{A_4^2}{A_5^3} - 2\frac{A_4}{A_5}\rho_{z_j x_1} + 2\frac{A_4 A_6}{A_5}\rho_{x_1 x_2}\right) \end{array}\right]. \qquad (5.3.1.7)$$

5.3.2 Proposed Estimator-VI

Motivated from Chand (1975) and Sousa et al. (2010), a ratio estimator under two-phase sampling based on scrambled response *y* have been proposed when observed response have positive correlation with non-sensitive variables x_1 and x_2. The proposed estimator t_{72} is given as,

$$t_{72} = \bar{z}_j'' \frac{\bar{x}_1'}{\bar{x}_1''} \frac{\bar{X}_2}{\bar{x}_2''}, \qquad \text{for } j = 4, G1, G2, G3, G4. \qquad (5.3.2.1)$$

To obtain the boas and mean square error of t_{72} in terms of e's

$$t_{72} = \bar{Z}_j\left(1 + e''_{\bar{z}_j}\right)\left(1 + e'_{\bar{x}_1}\right)\left(1 + e''_{\bar{x}_1}\right)^{-1}\left(1 + e''_{\bar{x}_2}\right)^{-1},$$

or

$$t_{72} - \bar{Z}_j = \bar{Z}_j\left[e''_{\bar{z}_j} + e'_{\bar{x}_1} - e''_{\bar{x}_1} - e'_{\bar{x}_1}e''_{\bar{x}_1} + e''^2_{\bar{x}_1} - e''_{\bar{x}_2} - e''_{\bar{z}_j}e''_{\bar{x}_2} - e'_{\bar{x}_1}e''_{\bar{x}_2} + e'^2_{\bar{x}_1}\right], \qquad (5.3.2.2)$$

Applying expectation on both sides, we may get the bias of t_{72} as,

$$Bias\left(t_{72}\right) = \bar{Z}_j\left[\theta_3\left(C_{x_1}^2 - C_{z_j x_1} + C_{x_1 x_2}\right) + \theta_2\left(C_{x_2}^2 - C_{z_j x_1}\right)\right]. \qquad (5.3.2.3)$$

Taking (5.3.2.2) up to first order of approximation and squaring on both sides we may obtain,

$$\left(t_{72}-\bar{Z}_j\right)^2=\bar{Z}_j^2\begin{bmatrix} e''^2_{\bar{z}_j}+e'^2_{\bar{x}_1}+e''^2_{\bar{x}_1}+e''^2_{\bar{x}_2}-2e'_{\bar{x}_1}e''_{\bar{z}_j}+2e''_{\bar{x}_1}e''_{\bar{z}_j} \\ +2e''_{\bar{x}_2}e''_{\bar{z}j}-2e'_{\bar{x}_1}e''_{\bar{x}_1}-2e''_{\bar{x}_1}e''_{\bar{x}_2}+2e'_{\bar{x}_1}e''_{\bar{x}_2} \end{bmatrix}. \quad (5.3.2.4)$$

Now taking expectation both sides we may have the *MSE* of t_{72} as,

$$MSE\left(t_{72}\right)=\bar{Z}_j^2\left[\theta_2\left(C_{z_j}^2+C_{x_2}^2+2C_{z_jx_2}\right)+\theta_3\left(C_{x_1}^2-2C_{z_jx_1}-2C_{x_1x_2}\right)\right]. \quad (5.3.2.5)$$

5.3.3 Proposed Estimator-VII

Motivated by Singh and Vishwakarma (2007), ratio-type exponential estimator under two-phase sampling have been proposed using two auxiliary variables based on observed response *z*. The estimator is as,

$$t_{73}=\bar{z}''_j\exp\left[\frac{\bar{x}'_1-\bar{x}''_1}{\bar{x}'_1+\bar{x}''_1}\right]\exp\left[\frac{\bar{X}_2-\bar{x}''_2}{\bar{X}_2+\bar{x}''_2}\right].$$

for j= 4, G1, G2, G3, G4 (5.3.3.1)

The estimator in (5.3.3.1) is expanded in terms of *e*'s and solving, we may have,

$$t_{73}-\bar{Z}_j=\bar{Z}_j\begin{bmatrix} e''_{\bar{z}_j}+\frac{e'_{\bar{x}_1}}{2}-\frac{e''_{\bar{x}_1}}{2}+\frac{e''^2_{\bar{x}_1}}{4}-\frac{e''_{\bar{x}_2}}{2}-\frac{e'_{\bar{x}_1}e''_{\bar{x}_2}}{4}+\frac{e''_{\bar{x}_1}e''_{\bar{x}_2}}{4} \\ +\frac{3e''^2_{\bar{x}_2}}{8}+\frac{e'_{\bar{x}_1}e''_{\bar{z}_j}}{2}-\frac{e''_{\bar{x}_1}e''_{\bar{z}_j}}{2}-\frac{e''_{\bar{x}_1}e''_{\bar{x}_2}}{2} \end{bmatrix}. \quad (5.3.3.2)$$

The expressions of the bias and *MSE* of t_{73} are obtained as,

$$Bias\left(t_{73}\right)=\bar{Z}_j\left[\frac{3}{8}\theta_3C_{x_1}^2-\frac{1}{2}\theta_3C_{z_jx_1}+\frac{3}{8}\theta_2C_{x_2}^2+\frac{1}{4}\theta_3C_{x_1x_2}-\frac{1}{2}\theta_2C_{z_jx_2}\right], \quad (5.3.3.3)$$

and

$$MSE\left(t_{73}\right)=\bar{Z}_j^2\left[\theta_2\left(C_{z_j}^2+\frac{1}{4}C_{x_2}^2-C_{z_jx_2}\right)+\theta_3\left(\frac{1}{4}C_{x_1}^2-C_{z_jx_1}+\frac{1}{2}C_{x_1x_2}\right)\right]. \quad (5.3.3.4)$$

5.3.4 Proposed Estimator-VIII

The generalized exponential-type proposed estimator have been proposed under partial-information case of two-phases sampling. The generalized exponential-type estimator is given by,

$$t_{74}^{(i)} = \left[\nu_0 \bar{z}_j'' + \nu_1\left(\bar{X}_2 - \bar{x}_2''\right)\right]\exp\left[\frac{a_2\left(\bar{x}_1' - \bar{x}_1''\right)}{\bar{x}_1{}' + \left(\nu_2 - 1\right)\bar{x}_1''}\right],$$

for j= 4, G1, G2, G3, G4. (5.3.4.1)

where $\nu_0 \neq 0$, ν_1, ν_2 are unknown constants whose values are to be estimated to obtain the minimum variance of $t_{74}^{(i)}$. Also, $a_2(0,-1,+1)$ is a generalization constant whose values provides regression estimator, class of regression-cum-exponential ratio and regression-cum-exponential product estimators respectively. However, for the different values of these constants we may have different class of estimators presented in Table 5.2.

To obtain the bias and the *MSE* of $t_{74}^{(i)}$, expand (5.3.4.1) in the form of e's the estimator $t_{74}^{(i)}$ becomes,

$$t_{74}^{(i)} = \left[\nu_0 \bar{Z}_j\left(1 + e_{\bar{z}}''\right) - \nu_1 \bar{X}_2 e_{\bar{x}_2}''\right]\exp\left[\frac{a_2}{\nu_2\left(1 + e_{\bar{x}_2}''\right)}\left(e_{\bar{x}_1}' - e_{\bar{x}_1}''\right)\left(1 + \frac{\left(e_{\bar{x}_1}' - e_{\bar{x}_1}''\right)}{\alpha_2\left(1 + e_{\bar{x}_1}''\right)}\right)\right],$$

(5.3.4.2)

Table 5.2
Class of Estimators for $t_{74}^{(i)}$

Class of estimators	ν_0	ν_1	a_1	ν_2
$t_{74}^{(1)}=\left(\nu_0\bar{z}_j''+\left(\bar{X}_2-\bar{x}_2''\right)\right)\left[\exp\left(\frac{\left(\bar{x}_1'-\bar{x}_1''\right)}{\bar{x}_1'}\right)\right]$	ν_0	1	1	1
$t_{74}^{(2)}=\left(\nu_0\bar{z}_j''+b_{z_jx_2}\left(\bar{X}_2-\bar{x}_2''\right)\right)\left[\exp\left(\frac{\left(\bar{x}_1'-\bar{x}_1''\right)}{\bar{x}_1''}\right)\right]$	ν_0	$b_{z_jx_2}$	1	1
$t_{74}^{(3)}=\left(\nu_0\bar{z}_j''+\rho_{z_jx_2}\left(\bar{X}_2-\bar{x}_2''\right)\right)\left[\exp\left(\frac{\left(\bar{x}_1'-\bar{x}_1''\right)}{\bar{x}_1''}\right)\right]$	ν_0	$\rho_{z_jx_2}$	1	1
$t_{74}^{(4)}=\left(\nu_0\bar{z}_j''+\left(\bar{X}_2-\bar{x}_2''\right)\right)\left[\exp\left(\frac{\bar{x}_1'-\bar{x}_1''}{2\left(\bar{x}_1'+\bar{x}_1''\right)}\right)\right]$	ν_0	1	½	2
$t_{74}^{(5)}=\left(\nu_0\bar{z}_j''+b_{z_jx_2}\left(\bar{X}_2-\bar{x}_2''\right)\right)\left[\exp\left(\frac{\bar{x}_1'-\bar{x}_1''}{2\bar{x}_1'}\right)\right]$	ν_0	$b_{z_jx_2}$	½	1
$t_{74}^{(6)}=\left(\nu_0\bar{z}_j''+C_{x_2}\left(\bar{X}_2-\bar{x}_2''\right)\right)\left[\exp\left(\frac{2\left(\bar{x}_1'-\bar{x}_1''\right)}{\bar{x}_1''}\right)\right]$	ν_0	C_{x2}	2	1
$t_{74}^{(7)}=\left(\nu_0\bar{z}_j''+\rho_{z_jx_1}\left(\bar{X}_2-\bar{x}_2''\right)\right)\left[\exp\left(\frac{\rho_{z_jx_2}\left(\bar{x}_1'-\bar{x}_1'\right)}{\bar{x}_1'}\right)\right]$	ν_0	$\rho_{z_jx_2}$	$\rho_{z_jx_1}$	1
$t_{74}^{(8)}=\left(\bar{z}_j''+b_{z_jx_2}\left(\bar{X}_2-\bar{x}_2''\right)\right)\left[\exp\left(\frac{\left(\bar{x}_1'-\bar{x}_1''\right)}{\bar{x}_1'}\right)\right]$	1	$b_{z_jx_2}$	1	1
$t_{74}^{(9)}=\left(\bar{z}_j''+\rho_{z_jx_2}\left(\bar{X}_2-\bar{x}_2''\right)\right)\left[\exp\left(\frac{\left(\bar{x}_1'-\bar{x}_1''\right)}{\bar{x}_1'}\right)\right]$	1	$\rho_{z_jx_2}$	1	1

Solving (5.3.4.2), we may have,

$$\begin{aligned} t_{74}^{(i)}-\bar{Z}_j &= \bar{Z}_j\left(\nu_0-1\right) \\ &\quad +\nu_0\bar{Z}_j\left[e''_{\bar{z}_j}+w\left(e'_{\bar{x}_1}-e''_{\bar{x}_1}\right)+\frac{1}{2}w^2\left(e'_{\bar{x}_1}-e''_{\bar{x}_1}\right)^2+w\left(e''_{\bar{z}_j}e'_{\bar{x}_1}-e''_{\bar{z}_j}e''_{\bar{x}_1}\right)\right] \\ &\quad -\nu_1\bar{X}_2e''_{\bar{x}_2}-\nu_1\bar{X}_2w\left(e''_{\bar{x}_2}e'_{\bar{x}_1}-e''_{\bar{x}_2}e''_{\bar{x}_1}\right), \end{aligned} \qquad (5.3.4.3)$$

where $w_2=a_2/\nu_2$. The expressions of the bias and *MSE* of estimator $t_{74}^{(i)}$ are obtained as,

$$Bias\left(t_{74}^{(i)}\right)=\bar{Z}_j\left(\nu_0-1\right)+\nu_0\bar{Z}_j\left[\frac{1}{2}\theta_3 w_2^2 C_{X1}^2-\theta_3 w_2 C_{z_j x_1}\right]+\nu_1\bar{X}_2\theta_3 w_2 C_{x_1x_2}, \tag{5.3.4.4}$$

and

$$MSE\left(t_{74}^{(i)}\right)=\bar{Z}_j^2\left(\nu_0-1\right)^2+\nu_0^2\bar{Z}_j^2\left[\theta_2 C_{z_j}^2+w_2^2\theta_3 C_{x_1}^2-2w_2\theta_3 C_{z_j x_1}\right]$$
$$+\nu_1^2\bar{X}_2^2\theta_2 C_{x_2}^2+2\bar{Z}_j^2\left(\nu_0-1\right)\nu_0\left[\frac{1}{2}w_2^2\theta_3 C_{X1}^2-w_2\theta_3 C_{z_j x_1}\right]$$
$$+2\nu_0\nu_1\bar{Z}_j\bar{X}_2\left[\theta_2 C_{z_j x_2}-\theta_3 w_2 C_{x_1x_2}\right]. \tag{5.3.4.5}$$

The *MSE* of $t_{74}^{(i)}$ is minimum for the optimum values w_2,ν_0 and ν_1 are obtained as,

$$w_2^{opt}=\frac{-\left(\bar{Z}_j C_{z_j}\rho_{z_j x_1}-\alpha_1 C_{x_2}\rho_{x_1x_2}\right)}{\theta_2\bar{Z}_j C_{z_j}},$$

$$\nu_1^{opt}=\frac{\nu_0\bar{Z}_j C_{z_j}\left(\theta_2\rho_{z_j x_2}-\theta_3\rho_{z_j x_1}\rho_{x_1x_2}\right)}{\bar{X}_2 C_{x_2}\left(\theta_3\rho_{x_1x_2}^2-\theta_2\right)}, \tag{5.3.4.6}$$

and

$$\nu_0^{opt}=\frac{1}{1+A_7 C_{z_j}^2}, \tag{5.3.4.7}$$

where

$$A_7=\left[\theta_3\rho_{z_j x_1}^2-\theta_2+\frac{\left(\theta_2\rho_{z_j x_2}-\theta_3\rho_{z_j x_1}\rho_{x_1x_2}\right)^2}{\left(\theta_3\rho_{x_1x_2}^2-\theta_2\right)}\right].$$

We may get the bias and *MSE*'s for the estimator $t_{74}^{(i)}$ (i=1,2,…,9) using different values of ν_0,ν_1,ν_2 and a_2 in (5.3.4.4-5.3.4.5).

Now, the minimum *MSE* of $t_{74}^{(i)}$ may be obtained using (5.3.4.6)-(5.3.4.7) in (5.3.4.5) as,

$$MSE\left(t_{74}^{(i)}\right)=\bar{Z}_j^2\left[1-\frac{1}{1+A_7 C_{z_j}^2}\right]. \tag{5.3.4.8}$$

The min MSE's for $t_{74}^{(i)}$ (i=1,2,…,9) can be obtained using different values of ν_0, ν_1, ν_2 and a_2 in (5.3.4.8).

5.3.5 Proposed Estimator-IX

Another proposed estimator is a ratio in regression estimator to estimate scrambled response y. The proposed ratio in regression estimator for the observed response z_j is given by,

$$t_{75} = \bar{z}_j'' + d_5 \left[\bar{x}_1' \frac{\bar{X}_2}{\bar{x}_2''} - \bar{x}_1'' \right] + d_6 \left(\bar{X}_2 - \bar{x}_2'' \right)$$

for j= 4, G1, G2, G3, G4 (5.3.5.1)

where d_5 and d_6 are suitable weights whose values provides minimum mean square error for estimator t_{75}.

Expanding (5.3.5.1) in terms of e's and solving, we may have,

$$t_{75} - \bar{Z}_j = \bar{Z}_j e''_{\bar{z}_j} + \bar{X}_1 d_5 \left[e'_{\bar{x}_1} - e''_{\bar{x}_2} - e'_{\bar{x}_1} e''_{\bar{x}_2} + e''^2_{\bar{x}_2} - e''_{\bar{x}_1} \right] - \bar{X}_2 d_6 e''_{\bar{x}_2} . \quad (5.3.5.2)$$

Taking expectation we may obtain the bias as,

$$Bias(t_{75}) = \bar{X}_1 d_5 C_{z_j} \left[\theta_2 C_{x_2} - \theta_1 C_{x_2} \rho_{x_1 x_2} \right]. \quad (5.3.5.3)$$

Rewrite (5.3.5.2) up to first order of approximation becomes,

$$t_{75} - \bar{Z}_j = \bar{Z}_j e''_{\bar{z}_j} + \bar{X}_1 d_5 \left[e'_{\bar{x}_1} - e''_{\bar{x}_2} - e''_{\bar{x}_1} \right] - \bar{X}_2 d_6 e''_{\bar{x}_2} . \quad (5.3.5.4)$$

Now, squaring and taking expectation above, the $MSE(t_{75})$ attained is as,

$$\begin{aligned} MSE(t_{75}) = {} & \bar{Z}_j^2 \theta_2 C_{z_j}^2 + d_5^2 \bar{X}_1^2 \left(\theta_2 C_{x_2}^2 + \theta_3 C_{x_1}^2 + 2\theta_3 C_{x_1 x_2} \right) \\ & + \theta_2 d_6^2 \bar{X}^2 C_{x_2}^2 - 2\bar{Z}\bar{X}_1 d_5 \left(\theta_3 C_{z_j x_1} + \theta_2 C_{z_j x_2} \right) \\ & - 2\bar{Z} d_6 \bar{X}_2 \theta_2 C_{z_j x_2} + 2\bar{X}_1 \bar{X}_2 d_5 d_6 \left(\theta_2 C_{X2}^2 + \theta_3 C_{x_1 x_2} \right). \end{aligned} \quad (5.3.5.5)$$

The $MSE\left(t_{75}\right)$ is minimum for the values of d_1 and d_2 are given as,

$$d_5^{opt}=\frac{\theta_2\bar{Z}_jC_{z_j}\left(B_1-B_2\rho_{z_jx_2}\right)}{\bar{X}_1\left(B_2+\theta_2B_3\right)} \text{ and } d_5^{opt}=\frac{\theta_2\bar{Z}_jC_{z_j}\rho_{z_jx_2}-\bar{X}_1d_5B_2}{\theta_2\bar{X}_2}, \tag{5.3.5.6}$$

where

$$B_1=\theta_3C_{x_1}\rho_{z_jx_1}+\theta_2C_{x_2}\rho_{z_jx_2},\ B_2=\theta_3C_{X1}\rho_{X1X2}+\theta_2C_{X2},$$

and $B_3=\theta_2C_{X2}^2+\theta_3C_{X1}^2+2\theta_3C_{X1X2}$.

So, $\min MSE\left(t_{75}\right)$ we may have,

$$\min MSE\left(t_{75}\right)=\bar{Z}_j^2\theta_2C_{z_j}^2\left[1+\rho_{z_jx_2}^2-\frac{\left(B_1-B_2\rho_{z_jx_2}\right)^2}{\left(B_2+\theta_2B_3\right)}\right]. \tag{5.3.5.7}$$

5.3.6 Proposed Estimator-X

Following Singh and Vishwakarma (2011) and Singh and Choudhary (2011), a different exponential estimator have been proposed for the case having two auxiliary non-sensitive variables is given by,

$$t_{76}=\bar{z}_j''\exp\left[\frac{\bar{x}_1'\dfrac{\bar{X}_2}{\bar{x}_2''}-\bar{x}_1''}{\bar{x}_1'\dfrac{\bar{X}_2}{\bar{x}_2''}+\bar{x}_1''}\right]\exp\left(\frac{\bar{X}_2-\bar{x}_2''}{\bar{X}_2+\bar{x}_2''}\right),$$

for j= 4, G1, G2, G3, G4. (5.3.6.1)

The estimator (5.3.6.1) is expanded in terms of e's and solving we may get,

$$\begin{aligned} t_{76}-\bar{Z}_j=\bar{Z}_j\Bigg[& e_{\bar{z}_j}''+\frac{1}{2}\left(e_{\bar{x}_1}'-e_{\bar{x}_2}''-e_{\bar{x}_1}''\right) \\ & +\frac{1}{8}\left(e_{\bar{x}_1}'^2+e_{\bar{x}_1}''^2+e_{\bar{x}_2}''^2-2e_{\bar{x}_1}'e_{\bar{x}_2}''-2e_{\bar{x}_1}'e_{\bar{x}_1}''+2e_{\bar{x}_1}''e_{\bar{x}_2}''\right)\frac{1}{2}e_{\bar{x}_2}''-\frac{1}{2}e_{\bar{z}_j}''e_{\bar{x}_2}'' \\ & +\frac{1}{8}e_{\bar{x}_2}''^2-\frac{1}{4}\left(e_{\bar{x}_1}'e_{\bar{x}_2}''-e_{\bar{x}_2}''^2-e_{\bar{x}_1}''e_{\bar{x}_2}''\right)-\frac{1}{2}\left(e_{\bar{x}_1}'e_{\bar{z}_j}''-e_{\bar{z}_j}''e_{\bar{x}_2}''-e_{\bar{z}_j}''e_{\bar{x}_1}''\right)\Bigg]. \end{aligned} \tag{5.3.6.2}$$

The bias and *MSE* expression of estimator t_{76} are obtained as,

$$Bias(t_{76}) = \frac{1}{2}\bar{Z}_j\left(\theta_2 C_{x_2}^2 - \theta_2 C_{z_j x_2} + \theta_3 C_{x_1 x_2} + \frac{1}{4}\theta_3 C_{x_1}^2 - \theta_3 C_{z_j x_1}\right), \quad (5.3.6.3)$$

and

$$MSE(t_{76}) = \bar{Z}_j^2\left[\theta_2\left(C_{z_j}^2 + C_{x_2}^2 - 2C_{z_j x_2}\right) + \theta_3\left(C_{X1}^2 - C_{z_j x_1} + C_{x_1 x_2}\right)\right]. \quad (5.3.6.4)$$

5.3.7 Proposed Estimator-XI

Adapting by Singh and Choudhary (2011), the ratio in exponential estimator has been proposed under two-phase sampling based on *SRR* models. The estimator t_{77} is presented as,

$$t_{77} = \bar{z}_j'' \exp\left[\frac{\bar{x}_1' \frac{\bar{X}_2}{\bar{x}_2''} - \bar{x}_1''}{\bar{x}_1' \frac{\bar{X}_2}{\bar{x}_2''} + \bar{x}_1''}\right], \qquad \text{for } j = 4, G1, G2, G3, G4. \quad (5.3.7.1)$$

Expand (5.3.7.1) in the form of *e*'s and solving, we may have,

$$t_{77} - \bar{Z}_j = \bar{Z}_j\left[e_{\bar{z}_j}'' + \frac{1}{2}\left(e_{\bar{x}_1}' - e_{\bar{x}_2}'' - e_{\bar{x}_1}''\right) + \frac{1}{2}\left(e_{\bar{x}_1}' e_{\bar{z}_j}'' - e_{\bar{z}_j}'' e_{\bar{x}_2}'' - e_{\bar{z}_j}'' e_{\bar{x}_1}''\right)\right.$$
$$\left. + \frac{1}{8}\left(e_{\bar{x}_1}'^2 + e_{\bar{x}_1}''^2 + e_{\bar{x}_2}''^2 - 2e_{\bar{x}_1}' e_{\bar{x}_2}'' - 2e_{\bar{x}_1}' e_{\bar{x}_1}'' + 2e_{\bar{x}_1}'' e_{\bar{x}_2}''\right)\right]. \quad (5.3.7.2)$$

The bias and *MSE* expression of estimator t_{77} are obtained as,

$$Bias(t_{77}) = \frac{1}{2}\bar{Z}_j\left(\frac{1}{4}\left(\theta_2 C_{x_2}^2 + \theta_3 C_{x_1}^2 + 2\theta_3 C_{x_1 x_2}\right) - \left(\theta_3 C_{z_j x_1} - \theta_2 C_{z_j x_2}\right)\right), \quad (5.3.7.3)$$

and

$$MSE(t_{77}) = \bar{Z}_j^2\left[\theta_2\left(C_{z_j}^2 + \frac{1}{4}C_{x_2}^2 - C_{z_j x_2}\right) + \theta_3\left(\frac{1}{4}C_{x_1}^2 - C_{z_j x_1} + \frac{1}{2}C_{x_1 x_2}\right)\right]. \quad (5.3.7.4)$$

5.3.8 Proposed Estimator-XII

The estimator t_{77} have been generalized for *PIC* two-phase sample and is given by,

$$t_{78}^{(i)} = \bar{z}_j'' \exp\left[\frac{c_2\left(\bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2''} - \bar{x}''\right)}{\bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2''} + (\omega_7 - 1)\bar{x}_1''}\right],$$

for *j*= 4, G1, G2, G3, G4. (5.3.8.1)

where c_2 and ω_7 are some suitably chosen constants whose values are to be estimated to obtain minimum variance of $t_{78}^{(i)}$. For the different choices of c_2 and ω_6, Table 5.3 provides some examples of different estimators.

Rewriting estimator $t_{78}^{(i)}$ in terms of *e*'s and by solving we get the expression as,

$$t_{78}^{(i)} - \bar{Z}_j = \bar{Z}_j \Big[e''_{\bar{z}_j} + c_3\left(e'_{\bar{x}_1} - e''_{\bar{x}_2} - e''_{\bar{x}_1}\right) + c_3\left(e'_{\bar{x}_1} e''_{\bar{z}_j} - e''_{\bar{z}_j} e''_{\bar{x}_2} - e''_{\bar{z}_j} e''_{\bar{x}_1}\right)$$
$$+ \frac{1}{2} c_3^2 \left(e'^2_{\bar{x}_1} + e''^2_{\bar{x}_1} + e''^2_{\bar{x}_2} - 2e'_{\bar{x}_1} e''_{\bar{x}_2} - 2e'_{\bar{x}_1} e''_{\bar{x}_1} + 2e''_{\bar{x}_1} e''_{\bar{x}_2}\right)\Big], \quad (5.3.8.2)$$

where $c_3 = c_2 / \omega_7$.

The expression of the bias of estimator $t_{78}^{(i)}$ obtained as,

$$Bias\left(t_{78}^{(i)}\right) = \bar{Z}_j \left(c_3^2 \frac{1}{2}\left(\theta_2 C_{x_2}^2 + \theta_3 C_{x_1}^2 + 2\theta_3 C_{x_1 x_2}\right) - c_3\left(\theta_3 C_{z_j x_1} + \theta_2 C_{z_j x_2}\right)\right).$$

(5.3.8.3)

To obtain mean square error of estimator $t_{78}^{(i)}$, we rewrite (5.3.8.3) up to first order of approximation, we have,

$$t_{78}^{(i)} - \bar{Z}_j = \bar{Z}_j \left(e''_{\bar{z}_j} + c_3\left(e'_{\bar{x}_1} - e''_{\bar{x}_2} - e''_{\bar{x}_1}\right)\right). \quad (5.3.8.4)$$

Table 5.3
Class of Estimators for $t_{78}^{(i)}$

Class of estimators	c_2	ω_7
$t_{78}^{(1)} = \bar{z}_j'' \exp\left[\dfrac{\left(\bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2''} - \bar{x}''\right)}{\bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2''}}\right]$	1	1
$t_{78}^{(2)} = \bar{z}_j'' \exp\left[\dfrac{\frac{1}{2}\left(\bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2''} - \bar{x}''\right)}{\bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2''} + \bar{x}_1''}\right]$	½	2
$t_{78}^{(2)} = \bar{z}_j'' \exp\left[\dfrac{\frac{1}{2}\left(\bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2''} - \bar{x}''\right)}{\bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2''}}\right]$	½	1
$t_{78}^{(4)} = \bar{z}_j'' \exp\left[\dfrac{2\left(\bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2''} - \bar{x}''\right)}{\bar{x}_1' \dfrac{\bar{X}_2}{\bar{x}_2''}}\right]$	2	1

The expression of *MSE* we may attain as,

$$MSE\left(t_{78}^{(i)}\right) = \bar{Z}_j^2 \left[\begin{array}{c} \theta_2 C_{z_j}^2 + c_3^2\left(\theta_3 C_{x_1}^2 + \theta_2 C_{x_2}^2 + 2\theta_3 C_{x_1x_2}\right) \\ -2c_3\left(\theta_3 C_{z_jx_1} + \theta_2 C_{z_jx_2}\right) \end{array}\right]. \qquad (5.3.8.5)$$

The optimum value of c_3 may be obtained as,

$$c_3^{opt} = \frac{\left(\theta_3 C_{z_jx_1} + \theta_2 C_{z_jx_2}\right)}{\left(\theta_3 C_{x_1}^2 + \theta_2 C_{x_2}^2 + \theta_3 C_{x_1x_2}\right)}. \qquad (5.3.8.6)$$

The expression of minimum *MSE* of generalized estimator $t_{78}^{(i)}$ obtained after substituting the value of c_3^{opt} as,

$$\min MSE\left(t_{78}^{(i)}\right)=\bar{Z}_j^2\left[\theta_2 C_{z_j}^2-\frac{\left(\theta_3 C_{z_j x_1}+\theta_2 C_{z_j x_2}\right)^2}{\left(\theta_3 C_{x_1}^2+\theta_2 C_{x_2}^2+\theta_3 C_{x_1 x_2}\right)}\right]. \qquad (5.3.8.7)$$

5.4 PROPOSED ESTIMATORS FOR NO-INFORMATION CASE (NIC)

In this section, the estimators presented in section 5.2. are modified for no-information case of two-phase sampling under *SRR* models. The proposed estimators are presented for the case when information on non-sensitive auxiliary variable are not available in advance of a survey in such a situation, it is usual to estimate them by sample mean $\bar{x}_1{}'$ and $\bar{x}_2{}'$ based on first-phase sample of size n_1 of which n_2 is a sub-sample $n_2 \subset n_1$.

5.4.1 Proposed Estimator-XIII

The regression estimator for *NIC* case using non-sensitive auxiliary variables is presented as,

$$t_{79}=\bar{z}_j''+b_5\left(\bar{x}_1{}'-\bar{x}_1''\right)+b_6\left(\bar{x}_2'-\bar{x}_2''\right),$$

for j= 4, G1, G2, G3, G4 (5.4.1.1)

where b_5 and b_6 are the sample regression coefficients between *Z* and X_1, and *Z* and X_2 respectively, scrambled on *y*. To the first order of approximation, the estimator t_{79} is unbiased, so to find the variance (5.4.1) may be expanded in terms of *e*'s, estimator t_{79} becomes,

$$t_{79}=\bar{Z}_j\left(1+e''_{\bar{z}_j}\right)+b_5\bar{X}_1\left(e'_{\bar{x}_1}-e''_{\bar{x}_1}\right)+b_6\bar{X}_2\left(e'_{\bar{x}_2}-e''_{\bar{x}_2}\right). \qquad (5.4.1.2)$$

Squaring on both sides we may get

$$\left(t_{79}-\bar{Z}_j\right)^2=\bar{Z}_j^2 e_{\bar{z}}''^2+b_5^2\bar{X}_1^2 e_{\bar{x}_1}'^2+b_5^2\bar{X}_1^2 e_{\bar{x}_1}''^2$$
$$-2b_5^2\bar{X}_1^2 e_{\bar{x}_1}' e_{\bar{x}_1}''+b_6^2\bar{X}_2^2 e_{\bar{x}_2}'^2+b_6^2\bar{X}_2^2 e_{\bar{x}_2}''^2-2b_6^2\bar{X}_2^2 e_{\bar{x}_2}' e_{\bar{x}_2}''$$
$$+2\bar{Z}_j\bar{X}_1 b_5 e_{\bar{z}_j}'' e_{\bar{x}_1}'-2\bar{Z}_j\bar{X}_1 b_5 e_{\bar{z}_j}'' e_{\bar{x}_1}''+2\bar{Z}_j\bar{X}_2 b_6 e_{\bar{z}_j}'' e_{\bar{x}_2}'$$
$$-2\bar{Z}_j\bar{X}_2 b_6 e_{\bar{z}_j}'' e_{\bar{x}_2}''+2\bar{X}_1\bar{X}_2 b_5 b_6 e_{\bar{x}_1}' e_{\bar{x}_2}'$$
$$-2\bar{X}_1\bar{X}_2 b_5 b_6 e_{\bar{x}_1}' e_{\bar{x}_2}''-2\bar{X}_1\bar{X}_2 b_5 b_6 e_{\bar{x}_1}'' e_{\bar{x}_2}'+2\bar{X}_1\bar{X}_2 b_5 b_6 e_{\bar{x}_1}'' e_{\bar{x}_2}'' . \quad (5.4.1.3)$$

Applying expectation both sides, we may obtain,

$$MSE\left(t_{79}\right)=\theta_2\bar{Z}_j^2 C_{z_j}^2+\theta_3\begin{bmatrix} b_5^2\bar{X}_1^2 C_{x_1}^2+b_6^2\bar{X}_2^2 C_{x_2}^2-2\bar{Z}_j\bar{X}_1 b_5 C_{z_j x_1} \\ -2\bar{Z}_j\bar{X}_2 b_6 C_{z_j x_2}+2\bar{X}_1\bar{X}_2 b_5 b_6 C_{x_1 x_2}\end{bmatrix}. \quad (5.4.1.4)$$

The $MSE\left(t_{79}\right)$ is minimum for the optimum values of sample regression coefficients b_5 and b_6 are as,

$$b_5^{opt}=\frac{\bar{Z}_j}{\bar{X}_1}\frac{C_{z_j}\left(\rho_{z_j x_1}-\rho_{z_j x_2}\rho_{x_1 x_2}\right)}{C_{x_1}\left(1-\rho_{x_1 x_2}^2\right)}=\frac{\bar{Z}}{\bar{X}_1}\frac{C_{z_j}}{C_{x_1}}\frac{A_8}{A_9}(Say), \quad (5.4.1.5)$$

and

$$b_6^{opt}=\frac{\bar{Z}_j}{\bar{X}_2}\frac{C_{z_j}\left(\rho_{z_j x_2}-\rho_{z_j x_1}\rho_{x_1 x_2}\right)}{C_{x_2}\left(1-\rho_{x_1 x_2}^2\right)}=\frac{\bar{Z}}{\bar{X}_2}\frac{C_{z_j}}{C_{x_2}}\frac{A_{10}}{A_9}(Say). \quad (5.4.1.6)$$

And the minimum value of $MSE\left(t_{79}\right)$ using optimum values (5.4.1.5)-(5.4.1.6) may be obtained as,

$$MSE\left(t_{79}\right)=\bar{Z}_j^2 C_{z_j}^2\begin{bmatrix}\theta_2+\frac{1}{A_9^2}\theta_3\left(A_8^2+A_{10}^2+2\rho_{x_1 x_2}A_9 A_{10}\right) \\ -2\frac{1}{A_9}\theta_3\left(A_8\rho_{z_j x_1}+A_{10}\rho_{z_j x_2}\right)\end{bmatrix}. \quad (5.4.1.7)$$

5.4.2 Proposed Estimator-XIV

The proposed ratio estimator under *NIC* two-phase sampling based on scrambled response *y* have is given as,

$$t_{80} = \bar{z}''_j \frac{\bar{x}_1{}'}{\bar{x}''_1}\frac{\bar{x}'_2}{\bar{x}''_2}, \qquad \text{for } j = 4, \text{G1, G2, G3, G4.} \tag{5.4.2.1}$$

To obtain the boas and mean square error of t_{80} in terms of *e*'s

$$t_{80} = \bar{Z}_j\left(1+e''_{\bar{z}_j}\right)\left(1+e'_{\bar{x}_1}\right)\left(1+e''_{\bar{x}_1}\right)^{-1}\left(1+e'_{\bar{x}_2}\right)\left(1+e''_{\bar{x}_2}\right)^{-1},$$

or

$$t_{80} - \bar{Z}_j = \bar{Z}_j\Big[e''_{\bar{z}_j} + e'_{\bar{x}_1} - e''_{\bar{x}_1} + e''_{\bar{z}_j}e'_{\bar{x}_1} - e''_{\bar{z}_j}e''_{\bar{x}_1} - e'_{\bar{x}_1}e''_{\bar{x}_1} + e''^2_{\bar{x}_1} + e'_{\bar{x}_2} + e''_{\bar{x}_2}$$
$$+ e''_{\bar{z}_j}e'_{\bar{x}_2} + e'_{\bar{x}_1}e'_{\bar{x}_2} - e'_{\bar{x}_1}e'_{\bar{x}_2} - e''_{\bar{z}_j}e''_{\bar{x}_2} + e'_{\bar{x}_1}e''_{\bar{x}_2} + e''_{\bar{x}_1}e''_{\bar{x}_2} - e'_{\bar{x}_2}e''_{\bar{x}_2} + e''^2_{\bar{x}_2}\Big]. \tag{5.4.2.2}$$

Applying expectation on both sides, we may get the bias of t_{80} as,

$$Bias\left(t_{80}\right) = \bar{Z}_j\theta_3\left[C^2_{x_2} + C_{x_1x_2} - C_{z_jx_1} - C_{z_jx_2}\right]. \tag{5.4.2.3}$$

Taking (5.4.9) up to first order of approximation and squaring on both sides we may obtain,

$$\left(t_{80} - \bar{Z}_j\right)^2 = \bar{Z}^2_j\Big[e''^2_{\bar{z}_j} + e'^2_{\bar{x}_1} + e''^2_{\bar{x}_1} + e'^2_{\bar{x}_2} + e''^2_{\bar{x}_2} + 2e'_{\bar{x}_1}e''_{\bar{z}_j} - 2e''_{\bar{x}_1}e''_{\bar{z}_j}$$
$$+ 2e''_{\bar{z}_j}e'_{\bar{x}_2} - 2e''_{\bar{x}_2}e''_{\bar{z}_j} - 2e'_{\bar{x}_1}e''_{\bar{x}_1} + 2e'_{\bar{x}_1}e'_{\bar{x}_2}$$
$$- 2e'_{\bar{x}_1}e''_{\bar{x}_2} - 2e'_{\bar{x}_2}e''_{\bar{x}_1} + 2e''_{\bar{x}_1}e''_{\bar{x}_2} - 2e'_{\bar{x}_2}e''_{\bar{x}_2}\Big]. \tag{5.4.2.4}$$

Now taking expectation both sides we may have the *MSE* of t_{80} as,

$$MSE\left(t_{80}\right) = \bar{Z}^2_j\left[\theta_2C^2_{z_j} + \theta_3\left(C^2_{x_1} + C^2_{x_2} - 2C_{z_jx_2} - 2C_{z_jx_1} + 2C_{x_1x_2}\right)\right]. \tag{5.4.2.5}$$

5.4.3 Proposed Estimator-XV

The exponential estimator have been proposed is given by,

$$t_{81} = \bar{z}''_j \exp\left[\frac{\bar{x}'_1 - \bar{x}''_1}{\bar{x}'_1 + \bar{x}''_1}\right]\exp\left[\frac{\bar{x}'_2 - \bar{x}''_2}{\bar{x}'_2 + \bar{x}''_2}\right],$$

for j= 4, G1, G2, G3, G4. (5.4.3.1)

To obtain the bias and the *MSE* of t_{81}, expand (5.4.3.1) in terms of e's and solving, we may have,

$$t_{81} - \bar{Z}_j = \bar{Z}_j\left[e''_{\bar{z}_j} + \frac{e'_{\bar{x}_1}}{2} - \frac{e''_{\bar{x}_1}}{2} + \frac{e'_{\bar{x}_2}}{2} - \frac{e''_{\bar{x}_2}}{2} + \frac{e'_{\bar{x}_1}e''_{\bar{z}_j}}{2} - \frac{e''_{\bar{x}_1}e''_{\bar{z}_j}}{2}\right.$$
$$+\frac{1}{8}\left(e'^2_{\bar{x}_1} + e''^2_{\bar{x}_1} - 2e'_{\bar{x}_1}e''_{\bar{x}_1}\right) + \frac{1}{8}\left(e'^2_{\bar{x}_2} + e''^2_{\bar{x}_2} - 2e'_{\bar{x}_2}e''_{\bar{x}_2}\right)$$
$$\left.+\frac{e'_{\bar{x}_1}e'_{\bar{x}_2}}{4} - \frac{e'_{\bar{x}_1}e''_{\bar{x}_2}}{4} - \frac{e''_{\bar{x}_1}e'_{\bar{x}_2}}{4} + \frac{e''_{\bar{x}_1}e''_{\bar{x}_2}}{4} + \frac{e'_{\bar{x}_2}e''_{\bar{z}_j}}{2} - \frac{e''_{\bar{x}_1}e''_{\bar{z}_j}}{2}\right]. \quad (5.4.3.2)$$

The expressions of the bias and *MSE* are obtained as,

$$Bias(t_{81}) = \frac{1}{2}\bar{Z}_j\theta_3\left[\frac{1}{4}\left(C^2_{x_1} + C^2_{x_2}\right) - \left(C_{z_jx_1} + C_{z_jx_2}\right) + \frac{1}{2}C_{x_1x_2}\right], \quad (5.4.3.3)$$

and

$$MSE(t_{81}) = \bar{Z}^2_j\left[\theta_2 C^2_{z_j} + \theta_3\left\{\frac{1}{4}\left(C^2_{x_1} + C^2_{x_2}\right) + \frac{1}{2}C_{x_1x_2} - \left(C_{z_jx_1} + C_{z_jx_2}\right)\right\}\right].$$

(5.4.3.4)

5.4.4 Proposed Estimator-XVI

The generalized exponential estimator for *NIC* is presented as,

$$t^{(i)}_{82} = \left[\kappa_0\bar{z}''_j + \kappa_1\left(\bar{x}'_2 - \bar{x}''_2\right)\right]\exp\left[\frac{a_3\left(\bar{x}'_1 - \bar{x}_1''\right)}{\bar{x}_1' + \left(\kappa_2 - 1\right)\bar{x}_1''}\right],$$

for j= 4, G1, G2, G3, G4, (5.4.4.1)

where $\kappa_0 \neq 0$, κ_1, κ_2 are unknown constants whose values are to be estimated to obtain the minimum variance of $t_{82}^{(i)}$. Also, $a_3(0,-1,+1)$ is a generalization constant whose values provides regression estimator, class of regression-cum-exponential ratio and regression-cum-exponential product estimators respectively. However, for the different values of these constants we may have different class of estimators presented in Table 5.4 below:

To obtain the bias and the *MSE* of $t_{82}^{(i)}$, (5.4.4.1) may be expanded in the form of *e*'s we may have the expression as,

$$t_{82}^{(i)} - \bar{Z}_j = \left[\kappa_0 \bar{Z}_j \left(1 + e''_{\bar{z}_j}\right) + \kappa_1 \bar{X}_2 \left(e'_{\bar{x}_2} - e''_{\bar{x}_2}\right)\right]$$
$$\exp\left[\frac{a_3\left(e'_{\bar{x}_2} - e''_{\bar{x}_2}\right)}{\kappa_2\left(1 + e''_{\bar{x}_1}\right)}\left(1 + \frac{\left(e'_{\bar{x}_2} - e''_{\bar{x}_2}\right)}{\kappa_2\left(1 + e''_{\bar{x}_1}\right)}\right)^{-1}\right], \qquad (5.4.4.2)$$

where $w_3 = a_3 / \kappa_2$.

Solving, we may obtain,

$$t_{82}^{(i)} - \bar{Z}_j = \bar{Z}_j\left(\kappa_0 - 1\right)$$
$$+\kappa_0 \bar{Z}_j\left[e''_{\bar{z}_j} + w_3\left(e'_{\bar{x}_1} - e''_{\bar{x}_1}\right) + \frac{1}{2}w_3^2\left(e'_{\bar{x}_1} - e''_{\bar{x}_1}\right)^2 + w_3\left(e'_{\bar{x}_1}e''_{\bar{z}_j} - e''_{\bar{x}_1}e''_{\bar{z}_j}\right)\right]$$
$$-\kappa_1 \bar{X}_2\left[\left(e'_{\bar{x}_2} - e''_{\bar{x}_2}\right) + w_3\left(e'_{\bar{x}_1}e'_{\bar{x}_2} - e'_{\bar{x}_1}e''_{\bar{x}_2} - e'_{\bar{x}_2}e''_{\bar{x}_2} + e''_{\bar{x}_1}e''_{\bar{x}_2}\right)\right]. \quad (5.4.4.3)$$

Table 5.4

Class of Estimators for $t_{82}^{(i)}$

Class of estimators	κ_0	κ_1	a_3	κ_2
$t_{82}^{(1)} = \left(\kappa_0 \bar{z}''_j + \left(\bar{x}'_2 - \bar{x}''_2\right)\right)\left[\exp\left(\frac{\left(\bar{x}'_1 - \bar{x}''_1\right)}{\bar{x}'_1}\right)\right]$	κ_0	1	1	1
$t_{82}^{(2)} = \left(\kappa_0 \bar{z}''_j + b_{z_j x_2}\left(\bar{x}'_2 - \bar{x}''_2\right)\right)\left[\exp\left(\frac{\left(\bar{x}'_1 - \bar{x}''_1\right)}{\bar{x}''_1}\right)\right]$	κ_0	$b_{z_j x_2}$	1	1
$t_{82}^{(3)} = \left(\kappa_0 \bar{z}''_j + \rho_{z_j x_2}\left(\bar{x}'_2 - \bar{x}''_2\right)\right)\left[\exp\left(\frac{\left(\bar{x}'_1 - \bar{x}''_1\right)}{\bar{x}'_1}\right)\right]$	κ_0	$\rho_{z_j x_2}$	1	1
$t_{82}^{(4)} = \left(\kappa_0 \bar{z}''_j + \left(\bar{x}'_2 - \bar{x}''_2\right)\right)\left[\exp\left(\frac{\bar{x}'_1 - \bar{x}''_1}{2\left(\bar{x}'_1 + \bar{x}''_1\right)}\right)\right]$	κ_0	1	½	2
$t_{82}^{(5)} = \left(\kappa_0 \bar{z}''_j + b_{z_j x_2}\left(\bar{x}'_2 - \bar{x}''_2\right)\right)\left[\exp\left(\frac{\bar{x}'_1 - \bar{x}''_1}{2\bar{x}'_1}\right)\right]$	κ_0	$b_{z_j x_2}$	½	1
$t_{82}^{(6)} = \left(\kappa_0 \bar{z}''_j + C_{X2}\left(\bar{x}'_2 - \bar{x}''_2\right)\right)\left[\exp\left(\frac{2\left(\bar{x}'_1 - \bar{x}''_1\right)}{\bar{x}'_1}\right)\right]$	κ_0	C_{x2}	2	1
$t_{82}^{(7)} = \left(\kappa_0 \bar{z}''_j + \rho_{z_j x_2}\left(\bar{x}'_2 - \bar{x}''_2\right)\right)\left[\exp\left(\frac{\rho_{z_j x_1}\left(\bar{x}'_1 - \bar{x}''_1\right)}{\bar{x}'_1}\right)\right]$	κ_0	$\rho_{z_j x_2}$	$\rho_{z_j x_1}$	1
$t_{82}^{(8)} = \left(\bar{z}''_j + b_{z_j x_2}\left(\bar{x}'_2 - \bar{x}''_2\right)\right)\left[\exp\left(\frac{\left(\bar{x}'_1 - \bar{x}''_1\right)}{\bar{x}'_1}\right)\right]$	1	$b_{z_j x_2}$	1	1
$t_{82}^{(9)} = \left(\bar{z}''_j + \rho_{z_j x_2}\left(\bar{x}'_2 - \bar{x}''_2\right)\right)\left[\exp\left(\frac{\left(\bar{x}'_1 - \bar{x}''_1\right)}{\bar{x}'_1}\right)\right]$	1	$\rho_{z_j x_2}$	1	1

The expressions of the bias and *MSE* are obtained as,

$$Bias\left(t_{82}^{(i)}\right) = \bar{Z}_j\left(\kappa_0 - 1\right) + \kappa_0 \bar{Z}_j \theta_3\left[\frac{1}{2} w_3^2 C_{x_1}^2 - w_3 C_{z_j} C_{x_1} \rho_{z_j x_1}\right]$$
$$+\kappa_1 \bar{X}_2 \theta_3 w_3 C_{x_1} C_{x_2} \rho_{x_1 x_2}, \qquad (5.4.4.4)$$

and

$$MSE\left(t_{82}^{(i)}\right)=\bar{Z}_j^2\left(\kappa_0-1\right)^2+\kappa_0^2\bar{Z}_j^2\left[\theta_2C_{z_j}^2+w_3^2\theta_3C_{x_1}^2-2w_3\theta_3C_{z_jx_1}\right]$$
$$+\kappa_1^2\bar{X}_2^2\theta_3C_{x_2}^2+2\bar{Z}^2\left(\kappa_0-1\right)\kappa_0\theta_3\left[\frac{1}{2}w_3^2C_{x_1}^2-w_3C_{z_jx_1}\right]$$
$$+2\left(\kappa_0-1\right)\kappa_1\bar{Z}_j\bar{X}_2\theta_3w_3C_{x_1x_2}+2\kappa_0\kappa_1\bar{Z}_j\bar{X}_2\theta_3\left(w_3C_{x_1x_2}-C_{z_jx_2}\right). \tag{5.4.4.5}$$

The *MSE* of $t_{82}^{(i)}$ is minimum for the optimum values of w_3,κ_0 and κ_1 are obtained as,

$$w_3^{opt}=\frac{C_{z_j}\left(\rho_{z_jx_1}-\kappa_0\rho_{z_jx_2}\rho_{x_1x_2}\right)}{C_{x_1}\left(1-\rho_{x_1x_2}^2\right)},\ \kappa_1^{opt}=\frac{\bar{Z}_jC_{z_j}\left(\kappa_0\rho_{z_jx_2}-\rho_{z_jx_1}\rho_{x_1x_2}\right)}{\bar{X}_2C_{x_2}\left(1-\rho_{x_1x_2}^2\right)} \tag{5.4.4.6}$$

and

$$\kappa_0^{opt}=\frac{1}{1+A_{11}C_{z_j}^2}, \tag{5.4.4.7}$$

where

$$A_{11}=\left[\theta_3\rho_{z_jx_2}^2+\theta_2+\frac{\left(\rho_{z_jx_2}-\rho_{z_jx_1}\rho_{x_1x_2}\right)^2}{\left(1-\rho_{x_1x_2}^2\right)}\right].$$

From (5.4.4.6-5.4.4.7), we may get the bias and *MSE*'s for $t_{82}^{(i)}$ (*i*=1,2,..,9) using different choices of $\kappa_0,\kappa_1,\kappa_2$ and a_3 in (5.4.4.4-5.4.4.5) as given in Table 5.4.

Now, the minimum *MSE* of $t_{82}^{(i)}$ may be obtained as,

$$MSE\left(t_{82}^{(i)}\right)=\bar{Z}_j^2\left[1-\frac{1}{1+A_{11}C_{z_j}^2}\right]. \tag{5.4.4.8}$$

We may obtain the min *MSE*'s for the estimator $t_{82}^{(i)}$ (*i*=1,2,..,9) using different choices of $\kappa_0,\kappa_1,\kappa_2$ and a_3 in (5.4.4.8).

5.4.5 Proposed Estimator-XVII:

The ratio in regression proposed estimator for the observed response *z* is given by,

$$t_{83} = \bar{z}''_j + d_7\left[\bar{x}'_1 \frac{\bar{x}'_2}{\bar{x}''_2} - \bar{x}''_1\right] + d_8\left(\bar{x}'_2 - \bar{x}''_2\right),$$

for *j*= 4, G1, G2, G3, G4. (5.4.5.1)

where d_7 and d_7 are suitable weights whose values provides minimum mean square error for estimator t_{83}.

Expanding (5.4.5.1) in terms of *e*'s and solving, we may have,

$$t_{83} - \bar{Z}_j = \bar{Z}_j e''_{\bar{z}_j} + \bar{X}_1 d_7\left[e'_{\bar{x}_1} + e'_{\bar{x}_2} - e''_{\bar{x}_2} - e''_{\bar{x}_1} + e'_{\bar{x}_1}e'_{\bar{x}_2} - e''_{\bar{x}_2}e'_{\bar{x}_1} + e''^2_{\bar{x}_2} - e'_{\bar{x}_2}e''_{\bar{x}_2}\right] + \bar{X}_2 d_8\left(e'_{\bar{x}_2} - e''_{\bar{x}_2}\right). \quad (5.4.5.2)$$

Taking expectation we may obtain the bias of t_{83} as,

$$Bias\left(t_{83}\right) = \bar{X}_1 d_7 C_{x_2}\left[\theta_2 C_{x_2} - \theta_1 C_{x_1}\rho_{x_1x_2}\right]. \quad (5.4.5.3)$$

Rewrite (5.4.5.2) up to first order of approximation becomes,

$$t_{83} - \bar{Z}_j = \bar{Z}_j e''_{\bar{z}_j} + \bar{X}_1 d_7\left[e'_{\bar{x}_1} + e'_{\bar{x}_2} - e''_{\bar{x}_2} - e''_{\bar{x}_1}\right] + \bar{X}_2 d_8\left(e'_{\bar{x}_2} - e''_{\bar{x}_2}\right). \quad (5.4.5.4)$$

Now, squaring and taking expectation above, the $MSE\left(t_{83}\right)$ attained is as,

$$\begin{aligned} MSE\left(t_{83}\right) = {} & \bar{Z}_j^2\theta_2 C_{z_j}^2 + d_7^2\bar{X}_1^2\theta_3\left(C_{x_2}^2 + C_{x_1}^2 + 2C_{x_1x_2}\right) \\ & + \theta_3 d_8^2\bar{X}_2^2 C_{x_2}^2 - 2\bar{Z}_j\bar{X}_1 d_7\theta_3\left(C_{z_jx_1} + C_{z_jx_2}\right) \\ & - 2\bar{Z}_j d_8\bar{X}_2\theta_3 C_{z_jx_2} + 2\bar{X}_1\bar{X}_2 d_7 d_8\theta_3\left(C_{x_2}^2 + C_{x_1x_2}\right). \end{aligned} \quad (5.4.5.5)$$

The $MSE\left(t_{83}\right)$ is minimum for the values of d_7 and d_8 are given as,

$$d_3^{opt} = \frac{\bar{Z}_j C_{z_j} \rho_{z_j x_2}}{\bar{X}_1 \left(C_{x_1} - C_{x_2} \rho_{x_1 x_2} \right)}, \text{ and } d_4^{opt} = \frac{\bar{Z}_j C_{Z_j} \rho_{z_j x_2}}{\bar{X}_2 C_{x_1} \rho_{x_1 x_2}}. \quad (5.4.5.6)$$

So, the $\min MSE\left(t_{83}\right)$ we may have,

$$\min MSE\left(t_{83}\right) = \bar{Z}_j^2 C_{z_j}^2 \left[\theta_2 + \theta_3 \left\{ \rho_{z_j x_1}^2 \frac{B_4}{B_5^2} - 2\rho_{z_j x_1} \frac{B_6}{B_5} + 2\rho_{z_j x_1} \rho_{z_j x_2} \frac{B_7}{B_5 \rho_{x_1 x_2}} + \frac{\rho_{z_j x_1}^2}{\rho_{x_1 x_2}^2} - 2 \frac{\rho_{z_j x_1}^2}{\rho_{x_1 x_2}} \right\} \right], \quad (5.4.5.7)$$

where

$B_4 = C_{x_1}^2 + C_{x_2}^2 + 2C_{x_1 x_2}$, $B_5 = C_{x_1} - C_{x_2} \rho_{x_1 x_2}$, $B_6 = C_{x_1} \rho_{z_j x_1} - C_{x_2} \rho_{z_j x_2}$,

and $B_7 = C_{x_2} + C_{x_1} \rho_{x_1 x_2}$.

5.4.6 Proposed Estimator-XVIII

The ratio in exponential estimator for two-phase no-information case is given by,

$$t_{84} = \bar{z}_j'' \exp \left[\frac{\bar{x}_1' \frac{\bar{x}_2'}{\bar{x}_2''} - \bar{x}_1''}{\bar{x}_1' \frac{\bar{x}_2'}{\bar{x}_2''} + \bar{x}_1''} \right]. \text{ for } j = 4, \text{G1, G2, G3, G4.} \quad (5.4.6.1)$$

To obtain the bias and the *MSE* of estimator in (5.4.6.1), we expand it in the form of *e*’s and solving, we may attain,

$$t_{84} - \bar{Z}_j = \bar{Z}_j \left[e_{\bar{z}_j}'' + \frac{1}{2}\left(e_{\bar{x}_1}' + e_{\bar{x}_2}' - e_{\bar{x}_2}'' - e_{\bar{x}_1}'' \right) + \frac{1}{2}\left(e_{\bar{x}_1}' e_{\bar{z}_j}'' + e_{\bar{x}_2}' e_{\bar{z}_j}'' - e_{\bar{z}_j}'' e_{\bar{x}_2}'' - e_{\bar{z}_j}'' e_{\bar{x}_1}'' \right) + \frac{1}{8}\begin{pmatrix} e_{\bar{x}_1}'^2 + e_{\bar{x}_2}'^2 + e_{\bar{x}_1}''^2 + e_{\bar{x}_2}''^2 + 2e_{\bar{x}_1}' e_{\bar{x}_2}' - 2e_{\bar{x}_1}' e_{\bar{x}_2}'' \\ -2e_{\bar{x}_1}' e_{\bar{x}_1}'' - 2e_{\bar{x}_2}' e_{\bar{x}_2}'' - 2e_{\bar{x}_2}' e_{\bar{x}_1}'' + 2e_{\bar{x}_1}'' e_{\bar{x}_2}'' \end{pmatrix} \right]. \quad (5.4.6.2)$$

The bias and *MSE* expression of estimator t_{84} are obtained as,

$$Bias(t_{84}) = \frac{1}{2}\bar{Z}_j\theta_3\left(\frac{1}{4}\left(C_{x_1}^2 + C_{x_2}^2 + 2C_{x_1x_2}\right) - \left(C_{z_jx_1} + C_{z_jx_2}\right)\right), \qquad (5.4.6.3)$$

and

$$MSE(t_{84}) = \bar{Z}_j^2\left[\theta_2 C_{z_j}^2 + \frac{1}{4}\theta_3\left(C_{x_1}^2 + C_{x_2}^2 + 2C_{x_1x_2}\right) - \theta_3\left(C_{z_jx_1} + C_{z_jx_2}\right)\right]. \qquad (5.4.6.4)$$

5.4.7 Proposed Estimator-XIX

The generalized exponential estimator for observed response z_j is presented by,

$$t_{85}^{(i)} = \bar{z}_j'' \exp\left[\frac{c_4\left(\bar{x}_1'\dfrac{\bar{x}_2'}{\bar{x}_2''} - \bar{x}_1''\right)}{\bar{x}_1'\dfrac{\bar{x}_2'}{\bar{x}_2''} + (\omega_8 - 1)\bar{x}_1''}\right].$$

for j= 4, G1, G2, G3, G4. (5.4.7.1)

Rewriting estimator $t_{85}^{(i)}$ in terms of e's and by solving we get the expression as,

$$\begin{aligned} t_{85}^{(i)} - \bar{Z}_j = \bar{Z}_j\Big[& e''_{\bar{z}_j} + c_4\left(e'_{\bar{x}_1} + e'_{\bar{x}_2} - e''_{\bar{x}_2} - e''_{\bar{x}_1}\right) \\ & + c_4\left(e'_{\bar{x}_1}e''_{\bar{z}} + e'_{\bar{x}_2}e''_{\bar{z}} - e''_{\bar{z}}e''_{\bar{x}_2} - e''_{\bar{z}}e''_{\bar{x}_1}\right) \\ & + \frac{1}{2}c_4^2\left(\begin{array}{l} e'^2_{\bar{x}_1} + e'^2_{\bar{x}_2} + e''^2_{\bar{x}_1} + e''^2_{\bar{x}_2} + 2e'_{\bar{x}_1}e'_{\bar{x}_2} - 2e'_{\bar{x}_1}e''_{\bar{x}_2} \\ \quad -2e'_{\bar{x}_1}e''_{\bar{x}_1} - 2e'_{\bar{x}_2}e''_{\bar{x}_1} - 2e'_{\bar{x}_2}e''_{\bar{x}_2} + 2e''_{\bar{x}_1}e''_{\bar{x}_2} \end{array}\right)\Big], \end{aligned} \qquad (5.4.7.2)$$

where $c_5 = c_4/\omega_8$.

The expression of bias of estimator $t_{85}^{(i)}$ obtained as,

$$Bias\left(t_{85}^{(i)}\right) = \theta_3\bar{Z}_j\left(c_5^2\frac{1}{2}\left(C_{x_2}^2 + C_{x_1}^2 + 2C_{x_1x_2}\right) - c_5\left(C_{z_jx_1} + C_{z_jx_2}\right)\right). \qquad (5.4.7.3)$$

To obtain mean square error of estimator $t_{85}^{(i)}$, we rewrite (5.4.7.2) up to first order of approximation, we have,

$$t_{85}^{(i)} - \bar{Z}_j = \bar{Z}_j\left(e''_{\bar{z}_j} + c_5\left(e'_{\bar{x}_1} + e'_{\bar{x}_2} - e''_{\bar{x}_2} - e''_{\bar{x}_1}\right)\right). \tag{5.4.7.4}$$

Taking square and expectation, the *MSE* we attain as,

$$MSE\left(t_{85}^{(i)}\right) = \bar{Z}_j^2\left[\theta_2 C_{z_j}^2 + c_5^2\theta_3\left(C_{x_1}^2 + C_{x_2}^2 + 2C_{x_1x_2}\right) - 2c_5\theta_3\left(C_{z_jx_1} + C_{z_jx_2}\right)\right]. \tag{5.4.7.5}$$

The optimum value of c_5 may be obtained as,

$$c_5^{opt} = \frac{\left(C_{z_jx_1} + C_{z_jx_2}\right)}{\left(C_{x_1}^2 + C_{x_2}^2 + 2C_{x_1x_2}\right)}. \tag{5.4.7.6}$$

The expression of minimum *MSE* obtained after substituting the value of c_5 as,

$$\min MSE\left(t_{85}^{(i)}\right) = \bar{Z}_j^2\left[\theta_2 C_{z_j}^2 - \frac{\left(C_{z_jx_1} + C_{z_jx_2}\right)^2}{\left(C_{x_1}^2 + C_{x_2}^2 + 2C_{x_1x_2}\right)}\right]. \tag{5.4.7.7}$$

5.4.8 Proposed Estimator XX

A different exponential estimator have been proposed for the case having two auxiliary non-sensitive variables is given by,

$$t_{86} = \bar{z}''_j \exp\left[\frac{\bar{x}'_1\frac{\bar{x}'_2}{\bar{x}''_2} - \bar{x}''_1}{\bar{x}'_1\frac{\bar{x}'_2}{\bar{x}''_2} + \bar{x}''_1}\right]\exp\left(\frac{\bar{x}'_2 - \bar{x}''_2}{\bar{x}'_2 + \bar{x}''_2}\right).$$

for j= 4, G1, G2, G3, G4. (5.4.8.1)

The estimator (5.3.43) is expanded in terms of e's and solving we may get,

$$t_{86}-\bar{Z}_j=\bar{Z}_j\Bigg[e''_{\bar{z}_j}+\frac{1}{2}\left(e'_{\bar{x}_1}+e'_{\bar{x}_2}-e''_{\bar{x}_2}-e''_{\bar{x}_1}\right)$$

$$+\frac{1}{8}\begin{pmatrix} e'^2_{\bar{x}_1}+e'^2_{\bar{x}_2}+e''^2_{\bar{x}_1}+e''^2_{\bar{x}_2}+2e'_{\bar{x}_1}e'_{\bar{x}_2}-2e'_{\bar{x}_1}e''_{\bar{x}_2} \\ -2e'_{\bar{x}_1}e''_{\bar{x}_1}-2e'_{\bar{x}_2}e''_{\bar{x}_2}-2e'_{\bar{x}_2}e''_{\bar{x}_2}+2e''_{\bar{x}_1}e''_{\bar{x}_2} \end{pmatrix}$$

$$+\frac{1}{2}e'_{\bar{x}_2}-\frac{1}{2}e''_{\bar{x}_2}+\frac{1}{2}e''_{\bar{z}_j}e'_{\bar{x}_2}-\frac{1}{2}e''_{\bar{z}_j}e''_{\bar{x}_2}+\frac{1}{8}\left(e'^2_{\bar{x}_2}+e''^2_{\bar{x}_2}-2e'_{\bar{x}_2}e''_{\bar{x}_2}\right)$$

$$+\frac{1}{2}\left(e'_{\bar{x}_1}e''_{\bar{z}_j}+e'_{\bar{x}_2}e''_{\bar{z}_j}-e''_{\bar{z}_j}e''_{\bar{x}_2}-e''_{\bar{z}_j}e''_{\bar{x}_1}\right)$$

$$+\frac{1}{4}\left(e'_{\bar{x}_1}e'_{\bar{x}_2}+e'^2_{\bar{x}_2}-e'_{\bar{x}_1}e''_{\bar{x}_2}-e'_{\bar{x}_2}e''_{\bar{x}_1}-e'_{\bar{x}_2}e''_{\bar{x}_2}+e''^2_{\bar{x}_2}-e'_{\bar{x}_2}e''_{\bar{x}_2}-e''_{\bar{x}_1}e''_{\bar{x}_2}\right)\Bigg]. \tag{5.4.8.2}$$

The bias and *MSE* expression of estimator t_{86} are obtained as,

$$Bias\left(t_{86}\right)=\frac{1}{2}\bar{Z}_j\theta_3\left[\frac{1}{4}\left(C^2_{x_1}+C^2_{x_2}+2C_{x_1x_2}\right)-\left(C_{z_jx_1}+C_{z_jx_2}\right)\right], \tag{5.4.8.3}$$

and

$$MSE\left(t_{86}\right)=\bar{Z}^2_j\Bigg[\theta_2C^2_{z_j}+\theta_3\Bigg\{\frac{1}{4}\left(C^2_{x_1}+2C^2_{x_2}+2C_{x_1x_2}\right).$$

$$+\frac{1}{2}\left(C^2_{x_2}+C_{x_1x_2}\right)-\left(C_{z_jx_1}+2C_{z_jx_2}\right)\Bigg\}\Bigg] \tag{5.4.8.4}$$

5.5 SIMULATION STUDY

In this section, the outcomes of the simulation study to investigate the performance of the proposed generalized estimators under additive model and proposed SRR models have been provided. In the simulation study, *MSE*'s of the estimators will be compared together empirically and theoretically. We assume two fixed populations size N=1000 generated from the multivariate normal distribution. The scrambling variables normally distributed as $S\sim N\left(0,10\%\sigma_{X_1}\right)$ and $R\sim N\left(1,10\%\sigma_{X_1}\right)$. The simulated multivariate normal distribution has the same theoretical means of $\left[Y,X_1,X_2\right]$ as $\mu=\left[5,5,5\right]$ and the different covariance matrices as given below:

Population I

$$\sigma^2 = \begin{bmatrix} 10 & 3 & 2.9 \\ 3 & 2 & 1.1 \\ 2.9 & 1.1 & 2 \end{bmatrix} \quad \rho_{X_1Y} = 0.6817, \quad \rho_{X_2Y} = 0.6705$$

Population II

$$\sigma^2 = \begin{bmatrix} 6 & 3 & 2.9 \\ 3 & 2 & 1.1 \\ 2.9 & 1.1 & 2 \end{bmatrix} \quad \rho_{X_1Y} = 0.8706, \quad \rho_{X_2Y} = 0.8428$$

For each population the first sample size considered is: n_1 = 50, 100, 200, 300 and the second sample size chosen are: n_2 = 20, 40, 80, 120.

For the simulation study, the *MSE'*s by means of 1000 samples of size n_1 and n_2 are selected from the population are estimated. Table (5.5-5.16) provides the comparison between the additive model and the proposed *SRR* models using non-sensitive auxiliary variable. In this study, the values for *a* and *g* used are 0.5 and 0.6 respectively. The empirical and theoretical results show us that as the coefficient of correlation i.e., ρ_{YX_1}, ρ_{YX_2} increases, the *MSE* of the estimators gets more and more effective.

5.6 DISCUSSION

In this chapter, we have proposed some estimators for two-phase sampling for three different cases regarding the availability of the auxiliary information. These three cases are defined as the full-information case (*FIC*), partial-information case (*PIC*) and no-information case (*NIC*). For case I, the *MSE* values of the estimators in (5.2.1.7), (5.2.2.6), (5.2.3.5) and (5.2.4.8) have been computed in Table (5.5-5.10). From Table (5.5-5.10), it is noticed that the proposed *SRR* models provides more efficient results for proposed estimators as compared to Pollock and Bek's (1971) model. Moreover, one can see that the proposed *SRR* model Z_{G1} obtain minimum *MSE*'s of the estimators. Similarly for case II, the *MSE* values of the estimators in (5.3.1.7), (5.3.2.5), (5.3.3.4), (5.3.4.8), (5.3.5.7), (5.3.6.4), (5.3.7.4) and (5.3.8.7) have been computed in Table (5.11-5.16). From Table (5.11-5.16) it is shown that as the sample size increases, the *MSE*'s decreases. It is also observed that the proposed estimators perform more efficiently under proposed *SRR* models as compared to Pollock and Bek's model. One can notice that the estimators t_{71},

$t_{74}^{(i)}$ and $t_{78}^{(i)}$ are more efficient among the estimators for case II. For case III, the *MSE* values of the estimators in (5.4.1.7), (5.4.2.5), (5.4.3.4), (5.4.4.8), (5.4.5.7), (5.4.6.4), (5.4.7.7) and (5.4.8.4) have been computed in Table (5.17-5.22). It is observed from the Table (5.17-5.22) that the proposed *SRR* models perform better than the additive model for all generalized proposed estimators.

Table 5.5

Simulation results at sample size n_1=50 and n_2=20 for the *MSE*s of the Estimators for Full-Information Case using Population I [N=1000, $\rho_{x1y} = 0.6817$ & $\rho_{x2y} = 0.6705$]

n_1=50 n_2=20		*MSE* Estimation using Population–I									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var\left(\bar{z}_j\right)$		0.54991	0.58153	0.47936	0.47777	0.48980	0.48421	0.48005	0.45320	0.53597	0.55390
Proposed Estimators	t_{67}	0.25509	0.27348	0.19392	0.20581	0.23725	0.23833	0.21442	0.22972	0.24480	0.24379
	t_{68}	0.26224	0.27511	0.20750	0.20621	0.24548	0.23968	0.21508	0.23793	0.25099	0.24469
	t_{69}	0.29784	0.33392	0.27302	0.26947	0.30533	0.30374	0.28118	0.29285	0.32423	0.30336
	$t_{70}^{(i)}$	0.24945	0.25976	0.19306	0.20375	0.23338	0.24591	0.20432	0.22483	0.24582	0.23647
	$t_{70}^{(1)}$	0.25245	0.26772	0.21105	0.21991	0.25200	0.26398	0.25452	0.23814	0.25838	0.24722
	$t_{70}^{(4)}$	0.27313	0.28664	0.20005	0.29514	0.25690	0.26984	0.21774	0.26883	0.27819	0.27925
	$t_{70}^{(7)}$	0.26306	0.25780	0.22417	0.26894	0.27407	0.29134	0.24971	0.25281	0.26310	0.25175
	$t_{70}^{(9)}$	0.26086	0.25106	0.20405	0.24693	0.23603	0.24885	0.21257	0.23921	0.25885	0.23276

Table 5.6

Simulation results at sample size n_1=100 and n_2=40 for the *MSE*s of the Estimators for Full-Information Case using Population I [N=1000, $\rho_{x1y} = 0.6817$ & $\rho_{x2y} = 0.6705$]

n_1=100 n_2=40		*MSE* Estimation using Population–I									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var\left(\bar{z}_j\right)$		0.26267	0.26303	0.23533	0.24001	0.24595	0.25158	0.24251	0.24072	0.25875	0.25170
Proposed Estimators	t_{67}	0.12430	0.12082	0.09026	0.09631	0.11610	0.11309	0.10399	0.09718	0.11809	0.11569
	t_{68}	0.12973	0.12122	0.09094	0.09744	0.11651	0.11397	0.10674	0.09769	0.11844	0.11706
	t_{69}	0.15512	0.15977	0.12785	0.12714	0.14662	0.14496	0.13699	0.12880	0.15271	0.15349
	$t_{70}^{(i)}$	0.12363	0.11294	0.08965	0.90013	0.11570	0.10301	0.10245	0.09176	0.10284	0.11372
	$t_{70}^{(1)}$	0.12843	0.12909	0.09966	0.09701	0.12489	0.11788	0.11282	0.10256	0.12405	0.12348
	$t_{70}^{(4)}$	0.14223	0.14187	0.09387	0.09134	0.12299	0.13349	0.12226	0.11195	0.13706	0.14111
	$t_{70}^{(7)}$	0.13387	0.13055	0.11911	0.09902	0.12169	0.12587	0.11564	0.10827	0.12487	0.13141
	$t_{70}^{(9)}$	0.12830	0.12716	0.10490	0.09841	0.12106	0.11776	0.10729	0.10128	0.11492	0.12305

Table 5.7
Simulation results at sample size n_1=300 and n_2=120 for the *MSE*s of the Estimators for Full-Information Case using Population I [N=1000, $\rho_{x1y} = 0.6817$ & $\rho_{x2y} = 0.6705$]

n_1=300 n_2=120		MSE Estimation using Population–I									
		Pollock and Bek Model		Proposed Models $g=0.6$ & $a=0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**
$Var\left(\bar{z}_j\right)$		0.07624	0.07585	0.06982	0.06999	0.07403	0.07388	0.07302	0.07101	0.07339	0.07247
Proposed Estimators	t_{67}	0.03668	0.03674	0.03185	0.03076	0.03307	0.03274	0.03338	0.03203	0.03424	0.03488
	t_{68}	0.03710	0.03688	0.03313	0.03081	0.03383	0.03284	0.03403	0.03213	0.03531	0.03502
	t_{69}	0.04498	0.04537	0.04371	0.03965	0.04322	0.04144	0.04438	0.04257	0.04192	0.04398
	$t_{70}^{(i)}$	0.03580	0.03531	0.03094	0.03006	0.03261	0.03320	0.03272	0.03065	0.03496	0.03005
	$t_{70}^{(1)}$	0.03787	0.03745	0.03345	0.03290	0.03518	0.03412	0.03423	0.03282	0.03383	0.03691
	$t_{70}^{(4)}$	0.04268	0.04343	0.03789	0.03625	0.04040	0.03991	0.03989	0.03715	0.03927	0.03831
	$t_{70}^{(7)}$	0.03933	0.03790	0.03604	0.03601	0.03698	0.03543	0.03692	0.03352	0.03628	0.03504
	$t_{70}^{(9)}$	0.03746	0.03620	0.03335	0.03268	0.03415	0.03383	0.03464	0.03213	0.03586	0.03620

Table 5.8

Simulation results at sample size n_1=50 and n_2=20 for the *MSE*s of the Estimators for Full-Information Case using Population II [N=1000, $\rho_{x1y} = 0.8706$ & $\rho_{x2y} = 0.8428$]

n_1=50 n_2=20		*MSE* Estimation using Population–II									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var\left(\bar{z}_j\right)$		0.36683	0.34114	0.30220	0.30783	0.31052	0.31761	0.31933	0.32475	0.32220	0.33297
Proposed Estimators	t_{67}	0.04650	0.04922	0.03660	0.04044	0.04121	0.04186	0.04109	0.04053	0.04517	0.04524
	t_{68}	0.04541	0.05001	0.03878	0.04071	0.04300	0.04192	0.04212	0.04097	0.04510	0.04736
	t_{69}	0.11187	0.11251	0.10214	0.09785	0.11054	0.01083	0.11144	0.10722	0.11047	0.10747
	$t_{70}^{(i)}$	0.04893	0.04994	0.04061	0.04120	0.04162	0.04072	0.04157	0.04031	0.04767	0.04133
	$t_{70}^{(1)}$	0.04560	0.04580	0.03763	0.04547	0.04583	0.04523	0.04491	0.04350	0.03832	0.04380
	$t_{70}^{(4)}$	0.09328	0.10720	0.08557	0.08589	0.08927	0.08318	0.08342	0.08197	0.08941	0.09251
	$t_{70}^{(7)}$	0.05025	0.06762	0.04147	0.04608	0.04507	0.04686	0.04660	0.04585	0.05014	0.04653
	$t_{70}^{(9)}$	0.04536	0.05794	0.03737	0.04141	0.04158	0.04273	0.04258	0.04248	0.03745	0.04172

Table 5.9

Simulation results at sample size n_1=100 and n_2=40 for the *MSE*s of the Estimators for Full-Information Case using Population II [N=1000, $\rho_{x1y} = 0.8706$ & $\rho_{x2y} = 0.8428$]

n_1=100 n_2=40		*MSE* Estimation using Population–II									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var\left(\bar{z}_j\right)$		0.15983	0.16644	0.13678	0.13655	0.14629	0.14972	0.14406	0.14660	0.15485	0.15616
Proposed Estimators	t_{67}	0.02484	0.02270	0.01096	0.01825	0.02178	0.02234	0.02034	0.02156	0.02239	0.02242
	t_{68}	0.02542	0.02394	0.01180	0.01864	0.02185	0.02264	0.02090	0.02193	0.02271	0.02286
	t_{69}	0.05434	0.05118	0.05064	0.04436	0.04855	0.05474	0.05398	0.05072	0.05450	0.05254
	$t_{70}^{(i)}$	0.02212	0.02428	0.00967	0.01826	0.21310	0.02214	0.02073	0.02173	0.02158	0.02203
	$t_{70}^{(1)}$	0.02500	0.02306	0.01721	0.01974	0.02422	0.02558	0.02375	0.02436	0.02410	0.02392
	$t_{70}^{(4)}$	0.04416	0.04540	0.04489	0.03701	0.04394	0.04846	0.04387	0.04694	0.04784	0.04361
	$t_{70}^{(7)}$	0.02649	0.02399	0.01681	0.02031	0.02434	0.02687	0.02412	0.02626	0.02370	0.02411
	$t_{70}^{(9)}$	0.02434	0.02288	0.01612	0.20316	0.02214	0.02377	0.02129	0.02038	0.02161	0.02459

Table 5.10
Simulation results at sample size n_1=300 and n_2=120 for the *MSE*s of the Estimators for Full-Information Case using Population II [N=1000, $\rho_{x1y} = 0.8706 \& \rho_{x2y} = 0.8428$]

n_1=300 n_2=120		*MSE* Estimation using Population–II									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var\left(\bar{z}_j\right)$		0.05045	0.06593	0.04419	0.04503	0.04663	0.04683	0.04632	0.04675	0.04933	0.04985
Proposed Estimators	t_{67}	0.00742	0.00764	0.00597	0.00617	0.00694	0.00668	0.00679	0.00674	0.00728	0.00733
	t_{68}	0.00776	0.00790	0.00634	0.00619	0.00685	0.00715	0.00692	0.00682	0.00729	0.00747
	t_{69}	0.01757	0.01578	0.05041	0.01644	0.01630	0.01505	0.01707	0.01705	0.01758	0.01659
	$t_{70}^{(i)}$	0.00715	0.00693	0.00556	0.00695	0.00674	0.00696	0.00683	0.00653	0.00713	0.00681
	$t_{70}^{(1)}$	0.01553	0.00783	0.00620	0.06223	0.00759	0.00752	0.00766	0.00747	0.01335	0.06448
	$t_{70}^{(4)}$	0.01401	0.01393	0.01326	0.01296	0.01333	0.01392	0.00770	0.00705	0.01885	0.01314
	$t_{70}^{(7)}$	0.00752	0.00861	0.00665	0.00716	0.00755	0.00752	0.00784	0.07134	0.07746	0.00808
	$t_{70}^{(9)}$	0.00747	0.00882	0.00604	0.00645	0.00687	0.00678	0.00703	0.07019	0.00735	0.00766

Table 5.11
Simulation results at sample size n_1=50 and n_2=20 for the *MSE*s of the Estimators for Partial-Information Case using Population I [N=1000, $\rho_{x1y} = 0.6817$ & $\rho_{x2y} = 0.6705$]

n_1=50 n_2=20		*MSE* Estimation using Population–I									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var(\bar{z}_j)$		0.54991	0.58153	0.47936	0.47777	0.48980	0.48421	0.48005	0.45320	0.53597	0.55390
Proposed Estimators	t_{71}	0.33981	0.33401	0.30548	0.30524	0.32003	0.32589	0.32745	0.32394	0.33684	0.32446
	t_{72}	0.26102	0.26303	0.20969	0.20888	0.24788	0.24391	0.24819	0.23028	0.25784	0.23448
	t_{73}	0.36890	0.36421	0.30138	0.30331	0.35424	0.34767	0.35962	0.35028	0.35028	0.34483
	$t_{74}^{(i)}$	0.04220	0.04228	0.03840	0.03894	0.03852	0.03074	0.04097	0.40212	0.04026	0.04085
	t_{75}	0.37665	0.37258	0.30174	0.30141	0.32781	0.32963	0.35206	0.35286	0.32472	0.32156
	t_{76}	0.26614	0.28841	0.23410	0.23450	0.25517	0.25710	0.24260	0.24931	0.25102	0.25261
	t_{77}	0.34782	0.34336	0.30275	0.30393	0.32963	0.33544	0.33459	0.33678	0.32472	0.31024
	$t_{78}^{(i)}$	0.05033	0.04948	0.04204	0.04292	0.04421	0.04422	0.03592	0.03432	0.04836	0.04821

Table 5.12

Simulation results at sample size n_1=100 and n_2=40 for the *MSE*s of the Estimators for Partial-Information Case using Population I [N=1000, $\rho_{x1y}=0.6817$ & $\rho_{x2y}=0.6705$]

n_1=100 n_2=40		*MSE* Estimation using Population–I									
		Pollock and Bek Model		Proposed Models $g=0.6$ & $a=0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var\left(\bar{z}_j\right)$		0.26267	0.26303	0.23533	0.24001	0.24595	0.25158	0.24251	0.24072	0.25875	0.25170
Proposed Estimators	t_{71}	0.14798	0.14967	0.12585	0.12440	0.13761	0.13246	0.13562	0.13967	0.13218	0.13570
	t_{72}	0.11386	0.11495	0.10288	0.10292	0.10401	0.10198	0.11187	0.11065	0.10146	0.10035
	t_{73}	0.16009	0.16591	0.13813	0.01345	0.01541	0.01504	0.15837	0.15892	0.15681	0.15789
	$t_{74}^{(i)}$	0.03868	0.03659	0.03404	0.03288	0.03693	0.03684	0.03707	0.03630	0.03276	0.03258
	t_{75}	0.16279	0.16249	0.13817	0.13743	0.15933	0.15270	0.14285	0.14150	0.15781	0.15289
	t_{76}	0.11513	0.13369	0.09471	0.09477	0.01093	0.01038	0.10656	0.02856	0.10204	0.10252
	t_{77}	0.13849	0.14118	0.10334	0.10345	0.11144	0.11533	0.11663	0.11819	0.12048	0.12553
	$t_{78}^{(i)}$	0.04223	0.04207	0.03510	0.03692	0.03851	0.03863	0.04032	0.04567	0.04021	0.04044

Table 5.13
Simulation results at sample size n_1=300 and n_2=120 for the *MSE*s of the Estimators for Partial-Information Case using Population I [N=1000, $\rho_{x1y}=0.6817$ & $\rho_{x2y}=0.6705$]

n_1=300 n_2=120		*MSE* Estimation using Population–I									
		Pollock and Bek Model		Proposed Models $g=0.6$ & $a=0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var(\bar{z}_j)$		0.07624	0.07585	0.06982	0.06999	0.07403	0.07388	0.07302	0.07101	0.07339	0.07247
Proposed Estimators	t_{71}	0.04795	0.05242	0.03494	0.03417	0.04684	0.04690	0.04110	0.04260	0.03468	0.03430
	t_{72}	0.04067	0.04116	0.03662	0.03611	0.03910	0.03954	0.03881	0.03937	0.03945	0.03964
	t_{73}	0.05196	0.05354	0.04398	0.04354	0.04550	0.04523	0.04544	0.04573	0.05176	0.05014
	$t_{74}^{(i)}$	0.01960	0.01895	0.01586	0.01584	0.01716	0.01718	0.01639	0.01654	0.01720	0.01707
	t_{75}	0.05724	0.05990	0.04315	0.04382	0.04600	0.04674	0.04504	0.04542	0.05431	0.05582
	t_{76}	0.04196	0.04165	0.03890	0.03877	0.04033	0.04046	0.03943	0.03990	0.04140	0.04056
	t_{77}	0.04722	0.04729	0.04048	0.04063	0.04347	0.04331	0.04341	0.04365	0.04505	0.04562
	$t_{78}^{(i)}$	0.03297	0.03586	0.02731	0.02731	0.02940	0.02930	0.02835	0.02985	0.03077	0.03089

Table 5.14

Simulation results at sample size n_1=50 and n_2=20 for the *MSE*s of the Estimators for Partial-Information Case using Population II [N=1000, $\rho_{x1y} = 0.8706$ & $\rho_{x2y} = 0.8428$]

n_1=50 n_2=20		***MSE* Estimation using Population–II**									
		Pollock and Bek Model		**Proposed Models** $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**	**Empirical**	**Theoretical**
$Var\left(\bar{z}_j\right)$		0.36683	0.34114	0.30220	0.30783	0.31933	0.32475	0.31052	0.31761	0.32220	0.33297
Proposed Estimators	t_{71}	0.08205	0.08419	0.07016	0.07151	0.07260	0.07480	0.07158	0.07200	0.07958	0.07742
	t_{72}	0.15259	0.16387	0.12948	0.12707	0.12602	0.13395	0.14019	0.13150	0.14399	0.14852
	t_{73}	0.18550	0.18606	0.15498	0.15236	0.16420	0.15396	0.16558	0.16017	0.16678	0.16455
	$t_{74}^{(i)}$	0.07160	0.07170	0.06053	0.06491	0.06689	0.06631	0.06545	0.06554	0.06669	0.06934
	t_{75}	0.22468	0.22699	0.20168	0.20026	0.21402	0.21834	0.21957	0.21507	0.21414	0.22053
	t_{76}	0.09856	0.09527	0.08040	0.08055	0.08459	0.08593	0.08707	0.08461	0.08858	0.08742
	t_{77}	0.01420	0.14819	0.11761	0.10019	0.12402	0.12847	0.12393	0.13165	0.13574	0.13572
	$t_{78}^{(i)}$	0.07563	0.07553	0.07045	0.07149	0.07305	0.07358	0.07298	0.07252	0.07342	0.07448

Table 5.15

Simulation result at sample size n_1=100 and n_2=40 for the *MSE*s of the Estimators for Partial-Information Case using Population II [N=1000, $\rho_{x1y} = 0.8706$ & $\rho_{x2y} = 0.8428$]

n_1=100 n_2=40		*MSE* Estimation using Population–II									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var\left(\bar{z}_j\right)$		0.15983	0.16644	0.13678	0.13655	0.14629	0.14972	0.14406	0.14660	0.15485	0.15616
Proposed Estimators	t_{71}	0.03686	0.03694	0.03275	0.03242	0.03458	0.03476	0.03381	0.03390	0.03526	0.03582
	t_{72}	0.07590	0.07593	0.07022	0.06275	0.07273	0.07284	0.07073	0.07158	0.07305	0.07372
	t_{73}	0.09727	0.09721	0.08993	0.07030	0.09556	0.09575	0.09088	0.09262	0.09637	0.09775
	$t_{74}^{(i)}$	0.03301	0.03343	0.03015	0.03054	0.03166	0.03236	0.03105	0.03156	0.03280	0.03290
	t_{75}	0.11733	0.11914	0.10458	0.10274	0.11004	0.11285	0.10564	0.10865	0.11212	0.11502
	t_{76}	0.04500	0.04901	0.04089	0.04061	0.04315	0.04256	0.04399	0.04392	0.04428	0.04489
	t_{77}	0.07426	0.07680	0.07130	0.07155	0.07314	0.07356	0.07278	0.07260	0.07366	0.07350
	$t_{78}^{(i)}$	0.03764	0.03871	0.02742	0.02610	0.02954	0.02959	0.02855	0.02865	0.03190	0.03274

Table 5.16
Simulation results at sample size n_1=300 and n_2=120 for the *MSE*s of the Estimators for Partial-Information Case using Population II [N=1000, $\rho_{x1y} = 0.8706$ & $\rho_{x2y} = 0.8428$]

n_1=300 n_2=120		*MSE* Estimation using Population–II									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var\left(\bar{z}_j\right)$		0.05045	0.06593	0.04419	0.04503	0.04663	0.04683	0.04632	0.04675	0.04933	0.04985
Proposed Estimators	t_{71}	0.01193	0.01133	0.01097	0.01062	0.01174	0.01185	0.01119	0.01113	0.01099	0.01095
	t_{72}	0.02323	0.02344	0.02074	0.02056	0.02274	0.02124	0.02109	0.02183	0.02209	0.02222
	t_{73}	0.03830	0.03706	0.02551	0.02440	0.02798	0.02960	0.02487	0.02682	0.03027	0.03187
	$t_{74}^{(i)}$	0.01057	0.01080	0.00805	0.00745	0.00911	0.00945	0.00829	0.00891	0.09703	0.09927
	t_{75}	0.04913	0.04864	0.03085	0.03084	0.03136	0.03255	0.03169	0.03166	0.03287	0.03370
	t_{76}	0.01525	0.01654	0.01019	0.01076	0.01277	0.01331	0.01127	0.01160	0.01459	0.01432
	t_{77}	0.02548	0.02617	0.01841	0.01864	0.02018	0.02351	0.01866	0.01920	0.02416	0.02459
	$t_{78}^{(i)}$	0.02421	0.02441	0.02262	0.01927	0.02147	0.02165	0.02046	0.02005	0.02208	0.02218

Table 5.17
Simulation results at sample size n_1=50 and n_2=20 for the *MSE*s of the Estimators for No-Information Case using Population I [N=1000, $\rho_{x1y} = 0.6817$ & $\rho_{x2y} = 0.6705$]

n_1=50 n_2=20		*MSE* Estimation using Population–I									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var\left(\bar{z}_j\right)$		0.54991	0.58153	0.47936	0.47777	0.48980	0.48421	0.48005	0.45320	0.53597	0.55390
Proposed Estimators	t_{79}	0.30723	0.30555	0.27336	0.27003	0.28703	0.28782	0.27283	0.28067	0.29282	0.29604
	t_{80}	0.31432	0.31542	0.27504	0.27377	0.28911	0.28923	0.27297	0.28381	0.29931	0.29895
	t_{81}	0.34422	0.34520	0.30977	0.30729	0.31587	0.32832	0.31479	0.31464	0.33134	0.34139
	$t_{82}^{(i)}$	0.28426	0.28644	0.26718	0.26387	0.26940	0.27520	0.26772	0.26653	0.27665	0.28179
	t_{83}	0.44200	0.44075	0.40331	0.40272	0.41407	0.42881	0.40860	0.40933	0.43914	0.43902
	t_{84}	0.35430	0.35740	0.32798	0.32771	0.33888	0.34010	0.33011	0.33027	0.34135	0.34952
	$t_{85}^{(i)}$	0.29436	0.29305	0.27394	0.27063	0.28518	0.28473	0.27970	0.27855	0.29051	0.29309
	t_{86}	0.34199	0.34086	0.30525	0.30634	0.32008	0.32685	0.31220	0.31976	0.33001	0.34007

Table 5.18

Simulation results at sample size n_1=100 and n_2=40 for the *MSE*s of the Estimators for No-Information Case using Population I [N=1000, $\rho_{x1y} = 0.6817$ & $\rho_{x2y} = 0.6705$]

n_1=100 n_2=40		*MSE* Estimation using Population–I									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var(\bar{z}_j)$		0.26267	0.26303	0.23533	0.24001	0.24595	0.25158	0.24251	0.24072	0.25875	0.25170
Proposed Estimators	t_{79}	0.16353	0.16989	0.14253	0.14501	0.15553	0.15926	0.15096	0.15023	0.16025	0.16092
	t_{80}	0.16408	0.17883	0.14498	0.14109	0.15605	0.15671	0.15924	0.15972	0.16095	0.16188
	t_{81}	0.19679	0.19402	0.16180	0.16369	0.18100	0.18092	0.17264	0.17408	0.18575	0.18846
	$t_{82}^{(i)}$	0.14636	0.14723	0.12071	0.12031	0.13064	0.13116	0.12313	0.12998	0.13608	0.13623
	t_{83}	0.23418	0.23842	0.20193	0.20168	0.21957	0.21900	0.21773	0.21715	0.22514	0.22946
	t_{84}	0.19682	0.19345	0.16183	0.16279	0.18102	0.18103	0.17734	0.17962	0.18577	0.18897
	$t_{85}^{(i)}$	0.15682	0.15848	0.13558	0.13115	0.14842	0.14836	0.13971	0.14010	0.15018	0.15203
	t_{86}	0.16978	0.16202	0.14911	0.14829	0.15450	0.15893	0.15378	0.15192	0.16562	0.16784

Table 5.19

Simulation results at sample size n_1=300 and n_2=120 for the *MSE*s of the Estimators for No-Information Case using Population I [N=1000, $\rho_{x1y}=0.6817$ & $\rho_{x2y}=0.6705$]

n_1=300 n_2=120		*MSE* Estimation using Population–I									
		Pollock and Bek Model		Proposed Models $g=0.6$ & $a=0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var\left(\bar{z}_j\right)$		0.07624	0.07585	0.06982	0.06999	0.07403	0.07388	0.07302	0.07101	0.07339	0.07247
Proposed Estimators	t_{79}	0.05117	0.05178	0.04027	0.04010	0.04694	0.04685	0.04528	0.04510	0.04715	0.04727
	t_{80}	0.05120	0.05179	0.04096	0.04084	0.04689	0.04689	0.04531	0.04583	0.04723	0.04757
	t_{81}	0.04806	0.04835	0.04365	0.04369	0.04783	0.04765	0.04406	0.04474	0.04776	0.04796
	$t_{82}^{(i)}$	0.03683	0.03698	0.03273	0.03107	0.35000	0.35028	0.03494	0.03426	0.35718	0.35768
	t_{83}	0.06732	0.06443	0.05366	0.05542	0.06189	0.06150	0.06071	0.06084	0.06274	0.06313
	t_{84}	0.04806	0.04854	0.04203	0.04266	0.04783	0.04707	0.04601	0.04638	0.04740	0.04765
	$t_{85}^{(i)}$	0.04118	0.04155	0.03890	0.03819	0.03947	0.03957	0.03953	0.03934	0.04033	0.04061
	t_{86}	0.04390	0.04519	0.04051	0.04057	0.04216	0.04225	0.04155	0.04134	0.04272	0.04311

Table 5.20

Simulation results at sample size n_1=50 and n_2=20 for the *MSE*s of the Estimators for No-Information Case using Population II [N=1000, $\rho_{x1y} = 0.8706$ & $\rho_{x2y} = 0.8428$]

n_1=50 n_2=20		*MSE* Estimation using Population–II									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var(\bar{z}_j)$		0.36683	0.34114	0.30220	0.30783	0.31052	0.31761	0.31933	0.32475	0.32220	0.33297
Proposed Estimators	t_{79}	0.19343	0.19023	0.12947	0.12897	0.15027	0.14540	0.13343	0.13010	0.16899	0.16244
	t_{80}	0.19659	0.19549	0.12482	0.12359	0.15135	0.14796	0.13557	0.13508	0.16947	0.16612
	t_{81}	0.21858	0.26119	0.14308	0.14277	0.19820	0.20510	0.17454	0.17630	0.19371	0.20595
	$t_{82}^{(i)}$	0.18262	0.19471	0.10959	0.10145	0.14231	0.15063	0.13337	0.14592	0.13702	0.15292
	t_{83}	0.31869	0.31690	0.23454	0.23538	0.29089	0.29521	0.27563	0.29339	0.29866	0.30404
	t_{84}	0.23758	0.25527	0.16443	0.16352	0.19834	0.19589	0.17465	0.17189	0.18380	0.18879
	$t_{85}^{(i)}$	0.15478	0.16906	0.12489	0.12678	0.15366	0.15353	0.13352	0.13486	0.13900	0.15209
	t_{86}	0.20547	0.19087	0.14496	0.14403	0.16776	0.16494	0.15105	0.16309	0.16935	0.18210

Table 5.21

Simulation results at sample size n_1=100 and n_2=40 for the MSEs of the Estimators for No-Information Case using Population II [N=1000, $\rho_{x1y} = 0.8706 \& \rho_{x2y} = 0.8428$]

n_1=100 n_2=40		MSE Estimation using Population–II									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var\left(\bar{z}_j\right)$		0.15983	0.16644	0.13678	0.13655	0.14406	0.14660	0.14629	0.14972	0.15485	0.15616
Proposed Estimators	t_{79}	0.08818	0.10571	0.05062	0.05035	0.06780	0.06537	0.06331	0.06456	0.07105	0.07530
	t_{80}	0.08905	0.11727	0.05133	0.05131	0.06878	0.06810	0.06353	0.06851	0.07194	0.07834
	t_{81}	0.10345	0.12828	0.07741	0.07822	0.084942	0.09137	0.08501	0.08609	0.98404	0.10635
	$t_{82}^{(i)}$	0.07000	0.07603	0.05690	0.05942	0.06380	0.06258	0.06006	0.06483	0.06737	0.07319
	t_{83}	0.18083	0.18206	0.011294	0.01161	0.13598	0.13017	0.12753	0.12483	0.13722	0.13812
	t_{84}	0.10882	0.10826	0.07743	0.08064	0.08496	0.09348	0.08452	0.09426	0.08406	0.09185
	$t_{85}^{(i)}$	0.07938	0.07988	0.06445	0.69320	0.07236	0.07546	0.06144	0.06203	0.07035	0.07546
	t_{86}	0.08387	0.09266	0.05664	0.05984	0.07326	0.07142	0.06966	0.07136	0.07461	0.07962

Table 5.22
Simulation results at sample size n_1=300 and n_2=120 for the *MSE*s of the Estimators for No-Information Case using Population II [N=1000, $\rho_{x1y} = 0.8706 \& \rho_{x2y} = 0.8428$]

n_1=300 n_2=120		*MSE* Estimation using Population–II									
		Pollock and Bek Model		Proposed Models $g = 0.6$ & $a = 0.5$							
				Z_{G1}		Z_{G2}		Z_{G3}		Z_{G4}	
Proposed Estimators		Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical	Empirical	Theoretical
$Var\left(\bar{z}_j\right)$		0.05045	0.06593	0.04419	0.04503	0.04683	0.04675	0.04663	0.04632	0.04933	0.04985
Proposed Estimators	t_{79}	0.04826	0.04723	0.01001	0.01046	0.02132	0.02469	0.01575	0.01439	0.03041	0.04561
	t_{80}	0.05553	0.05337	0.01086	0.01051	0.02146	0.02679	0.01590	0.01451	0.03076	0.04888
	t_{81}	0.07492	0.09274	0.02034	0.02039	0.02822	0.02842	0.02218	0.02200	0.03568	0.03865
	$t_{82}^{(i)}$	0.02438	0.02605	0.00529	0.00443	0.01965	0.01766	0.01461	0.01090	0.01905	0.02098
	t_{83}	0.05748	0.06610	0.03269	0.030588	0.040511	0.04059	0.03738	0.03098	0.04116	0.04688
	t_{84}	0.04612	0.04650	0.02225	0.02447	0.02822	0.02803	0.02218	0.02311	0.03568	0.04507
	$t_{85}^{(i)}$	0.02439	0.02952	0.00827	0.00872	0.02002	0.02152	0.01492	0.01563	0.02170	0.02373
	t_{86}	0.03644	0.03719	0.01976	0.02069	0.02104	0.02125	0.01809	0.02065	0.02262	0.02868

CHAPTER 6

PROPOSED ESTIMATORS TO ESTIMATE POPULATION VARIANCE BASED ON SCRAMBLED RANDOMIZED RESPONSE MODELS

6.1 INTRODUCTION

In this chapter, some estimators have been proposed to estimate population variance of the sensitive study variable based on additive scrambled model using non-sensitive auxiliary variables. The proposed estimators are presented in the form of observed response *z*. The expressions of the bias and mean square error have been derived up to first order of approximation.

6.2 PROPOSED VARIANCE ESTIMATORS FOR SINGLE-PHASE SAMPLING

6.2.1 Proposed Estimator-I

Adapting Bahl and Tuteja (1991), an exponential estimator for the observed response *z* is given by,

$$t_{87} = s_z^2 \exp\left[\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right]. \tag{6.2.1.1}$$

The expression in (6.2.1) may become by expanding in the form of *e*'s as,

$$t_{87} = S_z^2\left(1 + e_{s_z^2}\right)\exp\left[\frac{S_{x_1}^2 - S_{x_1}^2\left(1 + e_{s_{x_1}^2}\right)}{S_{x_1}^2 + S_{x_1}^2\left(1 + e_{s_{x_1}^2}\right)}\right],$$

or

$$t_{87} = S_z^2\left(1 + e_{s_z^2}\right)\exp\left[\frac{-e_{s_{x_1}^2}}{2}\left(1 + \frac{e_{s_{x_1}^2}}{2}\right)^{-1}\right]. \tag{6.2.1.2}$$

The bias and *MSE* expressions of estimator t_{s4} for first order of approximation are as,

$$Bias(t_{87}) = \theta_1 \frac{S_z^2}{2}\left(\frac{\vartheta_{02}}{2} - \vartheta_{22}\right), \tag{6.2.1.3}$$

and

$$MSE(t_{87}) = \theta S_z^4\left(\vartheta_{40} + \frac{\vartheta_{04}}{4} - \vartheta_{22}\right). \tag{6.2.1.4}$$

6.2.2 Proposed Estimator-II

Following Yadav and Kadilar (2013), a generalized exponential estimator to estimate population variance for sensitive variables is proposed as follows,

$$t_{88}^{(i)} = s_z^2 \exp\left[\frac{S_{x_1}^2 - s_{x_1}^2}{S_{x_1}^2 + (\alpha' - 1)s_{x_1}^2}\right] , \tag{6.2.2.1}$$

where α' is a positive real constant (i.e. $\alpha \geq 0$).

We may note that the choice of $\alpha' = 0$ in (6.2.2.1), $t_{88}^{(i)}$ reduces to

$$t_{88}^{(1)} = s_z^2 \exp(1), \tag{6.2.2.2}$$

which is biased and the *MSE* of $t_{88}^{(1)}$ is larger than s_z^2.

For $\alpha' = 1$ in (6.2.2.1), the $t_{88}^{(i)}$ reduces to

$$t_{88}^{(2)} = s_z^2 \exp\left[\frac{S_x^2 - s_x^2}{S_x^2}\right]. \tag{6.2.2.3}$$

For $\alpha' = 2$ in (6.2.2.4), the estimator $t_{88}^{(i)}$ may become

$$t_{88}^{(3)} = s_z^2 \exp\left[\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right] = t_{87}. \tag{6.2.2.4}$$

To obtain the bias and *MSE*, expanding (6.2.2.1) in terms of *e*'s, we have,

$$t_{88}^{(i)} = S_z^2\left(1+e_{s_z^2}\right)\exp\left[\frac{S_{x_1}^2 - S_{x_1}^2\left(1+e_{s_{x_1}^2}\right)}{S_{x_1}^2 + (\alpha'-1)S_{x_1}^2\left(1+e_{s_{x_1}^2}\right)}\right], \tag{6.2.2.5}$$

or

$$t_{88}^{(i)} = S_z^2\left(1+e_{s_z^2}\right)\exp\left[\frac{-e_{s_{x_1}^2}}{\alpha'}\left(1+\frac{(\alpha-1)e_{s_{x_1}^2}}{\alpha'}\right)^{-1}\right]. \tag{6.2.2.6}$$

We may obtain the expression of the bias and *MSE* of $t_{88}^{(i)}$ as,

$$Bias\left(t_{88}^{(i)}\right) = \theta_1 \frac{S_z^2}{2\alpha'^2}\vartheta_{04}\left[2\alpha'\left(1-\frac{\vartheta_{22}}{\vartheta_{04}}\right)-1\right], \tag{6.2.2.7}$$

and

$$MSE\left(t_{88}^{(i)}\right) = \theta_1 S_z^4\left[\vartheta_{40} + \frac{\vartheta_{04}}{\alpha'^2} - 2\frac{\vartheta_{22}}{\alpha'}\right]. \tag{6.2.2.8}$$

The $MSE\left(t_{88}^{(i)}\right)$ is minimum for the optimum value α' is as,

$$\alpha'^{opt} = \frac{\vartheta_{04}}{\vartheta_{22}}. \tag{6.2.2.9}$$

Substituting the optimum value α' in (6.2.2.8), we may get the $\min MSE\left(t_{88}^{(i)}\right)$ as follows,

$$\min MSE\left(t_{88}^{(i)}\right) = \theta_1 S_z^4\left[\vartheta_{40} - \frac{\vartheta_{22}^2}{\vartheta_{04}}\right]. \tag{6.2.2.9}$$

The estimator $t_{88}^{(i)}$ in (6.2.2.1) is to be used as an alternative to linear regression estimation in (2.2.16).

6.2.3 Proposed Estimator-III

Following Searl (1964) and Sanaullah et al. (2014) a generalized exponential estimator to estimate population variance of sensitive variable *y* is presented in the form of observed response *z* as

$$t_{89}^{(i)} = \eta s_z^2 \exp\left[\frac{S_{x_1}^2 - s_{x_1}^2}{S_{x_1}^2 + (\alpha' - 1) s_{x_1}^2}\right], \qquad (6.2.3.1)$$

where α' and $\eta > 0$ are some constants, assumed to be estimated such that the mean square error of $t_{89}^{(i)}$ is minimum.

However, for different values of these constants, we may get different estimators. The examples of different estimator are given in Table 6.1.

Expanding (6.2.12) in terms of *e*'s, we have,

$$t_{89}^{(i)} = \eta S_z^2 \left(1 + e_{s_z^2}\right) \exp\left[\frac{S_{x_1}^2 - S_{x_1}^2\left(1 + e_{s_{x_1}^2}\right)}{S_{x_1}^2 + (\alpha - 1) S_{x_1}^2 \left(1 + e_{s_{x_1}^2}\right)}\right],$$

or

$$t_{89}^{(i)} = (\eta - 1) S_z^2 + \eta S_z^2 \left(e_{s_z^2} - \frac{e_{s_{x_1}^2}}{\alpha} - \frac{e_{s_z^2} e_{s_{x_1}^2}}{\alpha} + \frac{e_{s_{x_1}^2}^2}{\alpha^2}\left(\alpha - \frac{1}{2}\right)\right). \qquad (6.2.3.2)$$

Table 6.1

Class of Estimator for $t_{89}^{(i)}$

Class of Estimators	η	α'
$t_{89}^{(1)} = \eta s_z^2 \exp(1)$	η	0
$t_{89}^{(2)} = \eta s_z^2 \exp\left[\frac{S_{x_1}^2 - s_{x_1}^2}{S_{x_1}^2}\right]$	η	1
$t_{89}^{(3)} = \eta s_z^2 \exp\left[\frac{S_{x_1}^2 - s_{x_1}^2}{S_{x_1}^2 + s_{x_1}^2}\right]$	η	2
$t_{89}^{(4)} = s_z^2 \exp\left[\frac{S_{x_1}^2 - s_{x_1}^2}{S_{x_1}^2 + (\alpha'-1) s_{x_1}^2}\right] = t_{88}^{(i)}$	1	α'
$t_{89}^{(5)} = s_z^2 \exp\left[\frac{S_{x_1}^2 - s_{x_1}^2}{S_{x_1}^2 + s_{x_1}^2}\right] = t_{87}$	1	2

We may obtain the expression of the bias and *MSE* of $t_{89}^{(i)}$ as,

$$Bias\left(t_{89}^{(i)}\right) = (\eta - 1) S_z^2 + \eta Bias\left(t_{88}^{(i)}\right), \tag{6.2.3.3}$$

and

$$MSE\left(t_{89}^{(i)}\right) = (\eta - 1)^2 S_z^4 + \eta^2 MSE\left(t_{88}^{(i)}\right) + 2\eta(\eta - 1) S_z^2 Bias\left(t_{88}^{(i)}\right). \tag{6.2.3.4}$$

The $MSE\left(t_{89}^{(i)}\right)$ is minimum for the optimum value η as,

$$\eta^{opt} = \frac{S_z^2\left[S_z^2 + Bias\left(t_{88}^{(i)}\right)\right]}{\left[S_z^4 + MSE\left(t_{88}^{(i)}\right) + 2S_z^2 Bias\left(t_{88}^{(i)}\right)\right]}. \tag{6.2.3.5}$$

We may get the bias and *MSE*'s for $t_{89}^{(i)}$ (*i*=1,2,..,5) using different values of η and α' in (6.2.3.3-6.2.3.4). Substituting the optimum value η in (6.2.3.5), we may get the $\min MSE\left(t_{89}^{(i)}\right)$ as follows,

$$\min MSE\left(t_{89}^{(i)}\right)=S_z^2\left(1-\frac{\left[1+\frac{Bias\left(t_{88}^{(i)}\right)}{S_z^2}\right]}{\left[1+\frac{MSE\left(t_{88}^{(i)}\right)}{S_z^4}+2\frac{Bias\left(t_{88}^{(i)}\right)}{S_z^2}\right]}\right). \tag{6.2.3.6}$$

From (6.2.3.4-6.2.3.6), it is clear that the $MSE\left(t_{89}^{(i)}\right)$ and $\min MSE\left(t_{89}^{(i)}\right)$ are the functions of $Bias\left(t_{88}^{(i)}\right)$ and $MSE\left(t_{88}^{(i)}\right)$. Also, one can get (6.2.3.4)-(6.2.3.5) directly if $Bias\left(t_{88}^{(i)}\right)$ and $MSE\left(t_{88}^{(i)}\right)$ are known.

6.3 PROPOSED VARIANCE ESTIMATORS FOR TWO-PHASE SAMPLING

In this section, the estimators presented in (6.2.1.1), (6.2.2.1) and (6.2.3.1) are modified for two-phase sampling no-information case (*NIC*) as below:

6.3.1 Proposed Estimator-IV

The exponential estimator for the observed response z under two-phase sampling is given as,

$$t_{90}=s_z''^2\exp\left[\frac{s_x'^2-s_x''^2}{s_x'^2+s_x''^2}\right]. \tag{6.3.1.1}$$

Expanding (6.3.1.1) in the form of e's we may have,

$$t_{90}=S_z^2\left(1+e''_{s_z^2}\right)\exp\left[\frac{S_{x_1}^2\left(1+e'_{s_{x_1}^2}\right)-S_{x_1}^2\left(1+e''_{s_{x_1}^2}\right)}{S_{x_1}^2\left(1+e'_{s_{x_1}^2}\right)+S_{x_1}^2\left(1+e''_{s_{x_1}^2}\right)}\right],$$

or

$$t_{90} = S_z^2\left(1+e''_{s_z^2}\right)\exp\left[\frac{e'_{s_{x_1}^2} - e''_{s_{x_1}^2}}{2}\left(1+\frac{e'_{s_{x_1}^2} + e''_{s_{x_1}^2}}{2}\right)^{-1}\right]. \tag{6.3.1.2}$$

The bias and *MSE* expressions of estimator t_{90} for first order of approximation are as,

$$Bias\left(t_{90}\right) = \theta_3\frac{S_z^2}{2}\left(\frac{\vartheta_{04}}{2}+\vartheta_{22}\right), \tag{6.3.1.3}$$

and

$$MSE\left(t_{90}\right) = S_z^4\left[\theta_2\vartheta_{40}+\theta_3\left(\frac{\vartheta_{04}}{4}-\vartheta_{22}\right)\right]. \tag{6.3.1.4}$$

6.3.2 Proposed Estimator-V

The generalized exponential estimator under two-phase sampling is given by,

$$t_{91}^{(i)} = s_z''^2\exp\left[\frac{s_{x_1}'^2 - s_{x_1}''^2}{s_{x_1}'^2+\left(\alpha''-1\right)s_{x_1}''^2}\right], \tag{6.3.2.1}$$

where α'' is a positive real constant (i.e. $\alpha'' \geq 0$). For different choices of α'', we may have different class of estimators for $t_{91}^{(i)}$ presented in Table 6.2.

To obtain the bias and *MSE*, expanding (6.3.2.1) in terms of *e*'s, we have,

$$t_{91}^{(i)} = S_z^2\left(1+e''_{s_z^2}\right)\exp\left[\frac{S_{x_1}^2\left(1+e'_{s_{x_1}^2}\right)-S_{x_1}^2\left(1+e''_{s_{x_1}^2}\right)}{S_{x_1}^2\left(1+e'_{s_{x_1}^2}\right)+\left(\alpha''-1\right)S_{x_1}^2\left(1+e''_{s_{x_1}^2}\right)}\right],$$

or

$$t_{91}^{(i)} = S_z^2\left(1+e''_{s_z^2}\right)\exp\left[\frac{1}{\alpha''}\left(e'_{s_{x_1}^2} - e''_{s_{x_1}^2}\right)^{-1}\right] \tag{6.3.2.2}$$

Table 6.2
Class of Estimator for $t_{91}^{(i)}$

Class of Estimators	α''
$t_{91}^{(1)} = s_z''^2 \exp(1)$	0
$t_{91}^{(2)} = s_z'^2 \exp\left[\dfrac{s_{x_1}'^2 - s_{x_1}''^2}{s_{x_1}'^2}\right]$	1
$t_{91}^{(3)} = s_z''^2 \exp\left[\dfrac{s_{x_1}'^2 - s_{x_1}''^2}{s_{x_1}'^2 + s_{x_1}''^2}\right]$	2
$t_{91}^{(4)} = s_z^2 \exp\left[\dfrac{S_{x_1}^2 - s_{x_1}^2}{S_{x_1}^2 + (\alpha - 1) s_{x_1}^2}\right]$	α''
$t_{91}^{(5)} = s_z^2 \exp\left[\dfrac{S_{x_1}^2 - s_{x_1}^2}{S_{x_1}^2 + s_{x_1}^2}\right]$	2

We may obtain the expression of the bias and *MSE* of $t_{91}^{(i)}$ as,

$$Bias\left(t_{91}^{(i)}\right) = \theta_3 \frac{S_z^2}{\alpha''}\left[\frac{1}{\alpha''}\vartheta_{04} - \vartheta_{22}\right], \tag{6.3.2.3}$$

and

$$MSE\left(t_{91}^{(i)}\right) = S_z^4\left[\theta_2 \vartheta_{40} + \theta_3 \frac{1}{\alpha''}\left(\frac{1}{\alpha''}\vartheta_{04} - 2\vartheta_{22}\right)\right]. \tag{6.3.2.4}$$

The $MSE\left(t_{91}^{(i)}\right)$ is minimum for the optimum value α'' as,

$$\alpha''^{opt} = \frac{\vartheta_{04}}{\vartheta_{22}}. \tag{6.3.2.5}$$

Substituting the optimum value α'' in (6.3.2.4), we may get the $\min MSE\left(t_{91}^{(i)}\right)$ as follows,

$$\min MSE\left(t_{91}^{(i)}\right) = S_z^4\left[\theta_2 \vartheta_{40} - \theta_3 \frac{\vartheta_{22}^2}{\vartheta_{04}}\right]. \tag{6.3.2.6}$$

We may get the bias and *MSE* for the estimator $t_{91}^{(i)}$ (i=1,2,..,5) using different values of α'' in (6.3.2.6) e.g. as presented in Table 6.2.

6.3.3 Proposed Estimator-VI

The generalized exponential estimator presented in (6.2.3.1) is modified for two-phase sampling to estimate population variance of sensitive variable *y* under the observed response is as

$$t_{92}^{(i)} = \eta' s_z''^2 \exp\left[\frac{s_{x_1}'^2 - s_{x_1}''^2}{s_{x_1}'^2 + (\alpha''-1) s_{x_1}''^2}\right], \tag{6.3.3.1}$$

where α'' and $\eta' > 0$ are some constants, assumed to be estimated such that the mean square error of $t_{92}^{(i)}$ is minimum. However, for different values of these constants, we may get different estimators. The examples of different estimator are given in Table 6.3.

Table 6.3

Class of Estimator for $t_{92}^{(i)}$

Class of Estimators	η'	α''
$t_{92}^{(1)} = \eta' s_z''^2 \exp(1)$	η'	0
$t_{92}^{(2)} = \eta' s_z''^2 \exp\left[\frac{s_{x_1}'^2 - s_{x_1}''^2}{s_{x_1}'^2}\right]$	η'	1
$t_{92}^{(3)} = \eta' s_z''^2 \exp\left[\frac{s_{x_1}'^2 - s_{x_1}''^2}{s_{x_1}'^2 + s_{x_1}''^2}\right]$	η'	2
$t_{92}^{(4)} = s_z''^2 \exp\left[\frac{s_{x_1}'^2 - s_{x_1}''^2}{s_{x_1}'^2 + (\alpha-1) s_{x_1}''^2}\right]$	1	α''
$t_{92}^{(5)} = s_z''^2 \exp\left[\frac{s_{x_1}'^2 - s_{x_1}''^2}{s_{x_1}'^2 + s_{x_1}''^2}\right]$	1	2

Expanding (6.3.3.1) in terms of e's, we have,

$$t_{92}^{(i)} = \eta' S_z^2 \left(1+e''_{s_z^2}\right) \exp\left[\frac{S_{x_1}^2\left(1+e'_{s_{x_1}^2}\right) - S_{x_1}^2\left(1+e''_{s_{x_1}^2}\right)}{S_{x_1}^2\left(1+e'_{s_{x_1}^2}\right) + \left(\alpha''-1\right) S_{x_1}^2\left(1+e''_{s_{x_1}^2}\right)}\right],$$

or

$$t_{92}^{(i)} - S_z^2 = \left(\eta'-1\right) S_z^2 + \eta' S_z^2 \left[e''_{s_z^2} + \frac{1}{\alpha''}\left(e'_{s_{x_1}^2} - e''_{s_{x_1}^2} + e''_{s_z^2} e'_{s_{x_1}^2} - e''_{s_z^2} e''_{s_{x_1}^2}\right) + \frac{1}{\alpha''^2}\left(e'^2_{s_{x_1}^2} + e''^2_{s_{x_1}^2} - 2e'_{s_{x_1}^2} e''_{s_{x_1}^2}\right)\right]. \quad (6.3.3.2)$$

The expressions of the bias and MSE of $t_{92}^{(i)}$ obtained are as,

$$Bias\left(t_{92}^{(i)}\right) = \left(\eta'-1\right) S_z^2 + \eta' Bias\left(t_{91}^{(i)}\right), \quad (6.3.3.3)$$

and

$$MSE\left(t_{92}^{(i)}\right) = \left(\eta'-1\right)^2 S_z^4 + \eta'^2 MSE\left(t_{91}^{(i)}\right) + 2\eta'\left(\eta'-1\right) S_z^2 Bias\left(t_{91}^{(i)}\right). \quad (6.3.3.4)$$

The MSE of $t_{92}^{(i)}$ is minimum for the optimum value η' as,

$$\eta'^{opt} = \frac{S_z^2\left[S_z^2 + Bias\left(t_{91}^{(i)}\right)\right]}{\left[S_z^4 + MSE\left(t_{91}^{(i)}\right) + 2S_z^2 Bias\left(t_{91}^{(i)}\right)\right]}. \quad (6.3.3.5)$$

We may get the bias and MSE for the estimator $t_{92}^{(i)}$ (i=1,2,..,5) using different values of η' and α'' in (6.1.16) as presented in Table 6.3. Substituting the optimum value of η' in (6.3.14), we may get the $\min MSE\left(t_{92}^{(i)}\right)$ as follows,

$$\min MSE\left(t_{92}^{(i)}\right)=S_z^2\left(1-\frac{\left[1+\dfrac{Bias\left(t_{91}^{(i)}\right)}{S_z^2}\right]}{\left[1+\dfrac{MSE\left(t_{91}^{(i)}\right)}{S_z^4}+2\dfrac{Bias\left(t_{91}^{(i)}\right)}{S_z^2}\right]}\right). \quad (6.3.3.6)$$

From (6.3.3.4-6.3.3.6), it is clear that the $MSE\left(t_{92}^{(i)}\right)$ and $\min MSE\left(t_{92}^{(i)}\right)$ are the functions of $Bias\left(t_{91}^{(i)}\right)$ and $MSE\left(t_{91}^{(i)}\right)$. Also, one can get (6.3.3.4-6.3.3.6) directly if $Bias\left(t_{91}^{(i)}\right)$ and $MSE\left(t_{91}^{(i)}\right)$ are known.

6.4 SIMULATION STUDY

In this section, we present results of the simulation study focusing on the performance of the estimators proposed under additive scrambled model. The following linear model is defined for the case where the sensitive variable *Y* and the auxiliary variable *X* are related to each other as:

$$Y_i = QX_i + e_i X_i^m$$

where $e_i \sim N(0,1)$ and the auxiliary variable $X_i \sim G(2.2,3.5)$ is generated from the gamma distribution. We consider a population of size *N*=3000 from the model for a given value of *m* and *Q*. From *SRSWOR*, we select the first-phase sample of sizes $n_1 =$80, 120, 160, 200 and 240 and the second-phase sample sizes are considered as $n_2 =$32, 48, 64, 80 and 96 from the given population of size *N*=3000.

Table 6.4
***MSE*'s of the Proposed Estimators for Single-Phase at *Q*=0.5 & *Q*=1.5**

Q	***n*₁**	**Estimator**		
		t_{87}	$t_{88}^{(i)}$	$t_{89}^{(i)}$
0.5	80	0.02204	**0.02069**	**0.01505**
	120	0.01530	**0.01439**	**0.00893**
	160	0.01057	**0.01044**	**0.00516**
	200	0.00811	**0.00748**	**0.00438**
	240	0.00680	**0.00632**	**0.00368**
1.5	80	0.04057	**0.04048**	**0.03875**
	120	0.02956	**0.02937**	**0.02019**
	160	0.01817	**0.01799**	**0.01649**
	200	0.01617	**0.01597**	**0.01452**
	240	0.01154	**0.01123**	**0.01070**

The scrambling variable '*S*' is generated from the beta distribution as $S \sim B(6.5, 0.5)$. The reported scrambled responses on study variable can be obtained by $Z_i = Y_i + S_i$, $i = 1, 2, ..., n$ from a given sample. The process will be repeated at *T*=5000 times.

In the simulation study, we fixed the values of *m*=0.5 and two values of *Q*=0.5 and 1.5. The *MSE* results of the proposed estimators discussed in section 6.2, have been presented in Table 6.4 for both values of *Q*. Table 6.5 presents the simulated results of proposed estimators presented in section 6.3.

6.5 DISCUSSION

The exponential-type estimators to estimate population variance of the sensitive variable under the observed response Z_4 have been proposed using auxiliary information and presented in this chapter. The bias and the *MSE* expressions have been derived for single-phase and two-phase sampling design. For each value of *Q*, the proposed estimators have been discussed separately. The simulated results of the *MSE* values of the estimators in

(6.2.1.4), (6.2.2.9) and (6.2.3.6) have been presented in Table 6.4. From Table 6.4, it is noticed that the proposed generalized estimators are more efficient than exponential estimator. Also, it is observed that the estimators proposed for single-phase sampling obtain minimum *MSE*'s at *Q*=1.5 as compared to *Q*=0.5. We conclude that $t_{89}^{(i)}$ is more efficient estimator than other proposed estimators. For two-phase sampling *NIC*, the *MSE* values of the proposed estimator in (6.3.1.4), (6.3.2.6) and (6.3.3.6) have been presented in Table 6.5. From Table 6.5, it is clear that the estimator $t_{92}^{(i)}$ is more efficient than $t_{90}^{(i)}$ and $t_{91}^{(i)}$. As compared to single phase estimators, it is clear that the value of *Q*=0.5 provides minimum *MSE* results for two-phase estimators than the simulated results of the proposed estimators at *Q*=1.5.

Table 6.5
***MSE*'s of the Proposed Estimators for Two-Phase Sampling at *Q*=0.5& *Q*=1.5.**

z	***n₁***	***n₂***	**Proposed Estimators**		
			t_{90}	$t_{91}^{(i)}$	$t_{92}^{(i)}$
0.5	80	32	0.10228	**0.10126**	**0.08251**
	120	48	0.06139	**0.06125**	**0.05966**
	160	64	0.04843	**0.04831**	**0.04323**
	200	80	0.04187	**0.04165**	**0.04006**
	240	96	0.03754	**0.03749**	**0.03204**
1.5	80	32	0.12719	**0.12702**	**0.11742**
	120	48	0.07331	**0.07278**	**0.06759**
	160	64	0.05373	**0.05358**	**0.04574**
	200	80	0.04182	**0.04170**	**0.03912**
	240	96	0.03307	**0.03285**	**0.02982**

REFERENCES

1. Asghar, A., Sanaullah, A. and Hanif, M. (2014). Generalized exponential type estimator for population variance in survey sampling. *Revista Colombiana de Estadística*, 37(1), 213-224.

2. Asghar, A., Sanaullah, A. and Hanif, M. (2017). A multivariate regression-cum-exponential estimator for population variance vector in two phase sampling. *Journal of King Saud University-Science*. (In Press).

3. Bahl, S. and Tuteja, R.K. (1991). Ratio and product type exponential estimator. Information and Optimization Sciences, 12, 159-163.

4. Bar-Lev, S.K., Bobovitch, E. and Boukai, B. (2004). A note on randomized response models for quantitative data. *Metrika*, 60(3), 255-260.

5. Bouza, C.N., Herrera, C. and Mitra, P.G. (2010). A review of randomized responses procedures: the qualitative variable case. *Rev InvOper*, 31(3), 240-247.

6. Bradburn, N.M., Sudman, S. and Wansink, B. (2004). *Asking questions: The definitive guide to questionnaire design for market research, political polls and social and health questionnaires*. San Francisco, CA: Jossey-Bass.

7. Chand, L. (1975). *Some ratio-type estimators based on two or more auxiliary variables*. Ph.D. Thesis. Iowa State Unviersity.

8. Chang, H.J. and Kuo, M.P. (2012). Estimation of population proportion in randomized response sampling using weighted confidence interval construction. *Metrika*, 75(5), 655-672.

9. Chaudhuri, A., Christofides, T.C. and Saha, A. (2009). Protection of privacy in efficient application of randomized response techniques. *Statistical Methods & Applications*, 18(3), 389-418.

10. Chen, C.C. and Singh, S. (2012). Estimation of multinomial proportions using higher order moments of scrambling variables in randomized response sampling. *Journal of Modern Applied Statistical Methods*, 11(1), 106-122.

11. Cochran, W.G. (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. *J. Agri. Science*, 30, 262-275.

12. Cochran, W.G. (1977). *Sampling Techniques*. New-York: John Wiley and Sons.

13. Diana, G. and Perri, P.F. (2010). New scrambled response models for estimating the mean of a sensitive quantitative character. *Journal of Applied Statistics*, 37(11), 1875-1890.

14. Diana, G. and Perri, P.F. (2011). A class of estimators for quantitative sensitive data. *Statistical Papers*, 52(3), 633-650.

15. Dihidar, K. and Chowdhury, J. (2013). Enhancing a randomized response model to estimate population means to sensitive questions. *Mathematical Population Studies*, 20(3), 123-136.

16. Eichhorn, B.H. and Hayre, L.S. (1983). Scrambled randomized response methods for obtaining sensitive quantitative data. *Journal of Statistical Planning and Inference*, 7(4), 307-316.

17. Gjestvang, C.R. and Singh, S. (2009). An improved randomized response model: estimation of mean. *Journal of Applied Statistics*, 36(12), 1361-1367.

18. Greenberg, B.G., Abul-Ela, A.L.A., Simmons, W.R. and Horvitz, D.G. (1969). The unrelated question randomized response model: Theoretical framework. *Journal of the American Statistical Association*, 64(326), 520-539.

19. Greenberg, B.G., Kuebler Jr, R.R., Abernathy, J.R. and Horvitz, D.G. (1971). Application of the randomized response technique in obtaining quantitative data. *Journal of the American Statistical Association*, 66(334), 243-250.

20. Grover, L.K. and Kaur, P. (2011). An improved exponential estimator of finite population mean in simple random sampling using an auxiliary attribute. *Applied Mathematics and Computation*, 218(7), 3093-3099.

21. Grover, L.K. and Kaur, P. (2014). A generalized class of ratio type exponential estimators of population mean under linear transformation of auxiliary variable. *Communications in Statistics-Simulation and Computation*, 43(7), 1552-1574.

22. Guerriero, M. and Sandri, M.F. (2007). A note on the comparison of some randomized response procedures. *Journal of Statistical Planning and Inference*, 137(7), 2184-2190.

23. Gupta, S. and Shabbir, J. (2004). Sensitivity estimation for personal interview survey questions. *Statistica*, 64(4), 643-653.

24. Gupta, S. and Shabbir, J. (2008). On improvement in estimating the population mean in simple random sampling. *Journal of Applied Statistics*, 35(5), 559-566.

25. Gupta, S., Gupta, B. and Singh, S. (2002). Estimation of sensitivity level of personal interview survey questions. *Journal of Statistical Planning and Inference*, 100(2), 239-247.

26. Gupta, S., Kalucha, G. and Shabbir, J. (2017). A regression estimator for finite population mean of a sensitive variable using an optional randomized response model. *Communications in Statistics-Simulation and Computation*, 46(3), 2393-2405.

27. Gupta, S., Shabbir, J. and Sehra, S. (2010). Mean and sensitivity estimation in optional randomized response models. *Journal of Statistical Planning and Inference*, 140(10), 2870-2874.

28. Gupta, S., Shabbir, J., Sousa, R. and Corte-Real, P. (2012). Estimation of the mean of a sensitive variable in the presence of auxiliary information. *Communications in Statistics-Theory and Methods*, 41(13-14), 2394-2404.

29. Gupta, S., Shabbir, J., Sousa, R. and Corte-Real, P. (2016). Improved Exponential Type Estimators of the Mean of a Sensitive Variable in the Presence of Non-sensitive Auxiliary Information. *Communications in Statistics-Simulation and Computation*, 45(9), 3317-3328.

30. Himmelfarb, S. and Edgell, S.E. (1980). Additive constants model: A randomized response technique for eliminating evasiveness to quantitative response questions. *Psychological Bulletin*, 87(3), 525-530.

31. Huang, K.C. (2010). Unbiased estimators of mean, variance and sensitivity level for quantitative characteristics in finite population sampling. *Metrika*, 71(3), 341-352.

32. Hussain, Z. (2012). Improvement of the Gupta and Thornton scrambling model through double use of randomization device. *International Journal of Academic Research in Business and Social Sciences*, 2(6), 91-97.

33. Hussain, Z., Al-Sobhi, M.M. and Al-Zahrani, B. (2014). Additive and subtractive scrambling in optional randomized response modeling. *Plos One*, 9(1), e83557.

34. Hussain, Z. and Al-Zahrani, B. (2016). Mean and sensitivity estimation of a sensitive variable through additive scrambling. *Communications in Statistics-Theory and Methods*, 45(1), 182-193.

35. Hussain, Z. and Khan, K. (2013). On estimation of sensitive mean using scrambled data. *World Applied Sciences Journal*, 23(9), 1201-1206.

36. Isaki, C.T. (1983). Variance estimation using auxiliary information. *Journal of the American Statistical Association*, 78(381), 117-123.

37. Kadilar, C. and Cingi, H. (2006). Ratio estimators for the population variance in simple and stratified random sampling. *Applied Mathematics and Computation*, 173(2), 1047-1059.

38. Kalucha, G., Gupta, S. and Dass, B.K. (2015). Ratio estimation of finite population mean using optional randomized response models. *Journal of Statistical Theory and Practice*, 9(3), 633-645.

39. Khan, M. (2016). A ratio chain-type exponential estimator for finite population mean using double sampling. *Springer Plus*, 5(1), 1.

40. Khan, M. and Shabbir, J. (2013). A Ratio Type Estimator for the Estimation of Population Variance using Quartiles of an Auxiliary Variable. *Journal of Statistics Applications & Probability*, 2(3), 319-325.

41. Khare, B.B., Srivastava, U. and Kumar, K. (2013). A generalized chain ratio in regression estimator for population mean using two auxiliary characters in sample survey. *J Sci Res Banaras Hindu Univ Varanasi*, 57, 147-153.

42. Khoshnevisan, M., Singh, R., Chauhan, P., Sawan, N. and Smarandache, F. (2007). A general family of estimators for estimating population mean using known value of some population parameter(s). *Far East Journal of Theoretical Statistics*, 22, 181-191.

43. Khoshnevisan, M., Singh, H.P., Singh, S. and Smarandache, F. (2002). A general class of estimators of population median using two auxiliary variables in double sampling. *Randomness and Optimal Estimation in Data Sampling*, 26-39.

44. Kiregyera, B. (1984). Regression-type estimators using two auxiliary variables and the model of double sampling from finite populations. *Metrika*, 31(1), 215-226.

45. Koyuncu, N. and Kadilar, C. (2009). Family of estimators of population mean using two auxiliary variables in stratified random sampling. *Commun. Stat., Theory Methods*, 38(14), 2398-2417.

46. Koyuncu, N., Gupta, S. and Sousa, R. (2014). Exponential-type estimators of the mean of a sensitive variable in the presence of non-sensitive auxiliary information. *Communications in Statistics-Simulation and Computation*, 43(7), 1583-1594.

47. Laplace, P.S. (1951). *A philosophical essay on probabilities*. English translation of the French original from 1820 by FW Truscott and FL Emory.

48. Mehta, S., Dass, B.K., Shabbir, J. and Gupta, S. (2012). A three-stage optional randomized response model. *Journal of Statistical Theory and Practice*, 6(3), 417-427.

49. Murthy, M.N. (1964). Product method of estimation. *Sankhya*, 26, 294-307.

50. Mushtaq, N. and Noor-ul-Amin, M. (2016). Estimation of a Population Mean of a Sensitive Variable in Stratified Two-Phase Sampling. *Pak. J. Statist*, 32(5), 393-404.

51. Neyman, J. (1934). On the two different aspects of representative method: The method of stratified sampling and the method of purposive selection. *J. Roy. Statist. Soc.*, 97, 558-606.

52. Neyman, J. (1938). Contribution to the theory of sampling human populations. *Journal of the American Statistical Association*, 33(201), 101-116.

53. Pollock, K.H. and Bek, Y. (1976). A comparison of three randomized response models for quantitative data. *Journal of the American Statistical Association*, 71(356), 884-886.

54. Prasad, B. (1989). Some improved ratio type estimators of population mean and ratio in finite population sample surveys. *Communications in Statistics-Theory and Methods*, 18(1), 379-392.

55. Rao. T.J. (1991). On certain methods of improving ratio and regression estimators, *Communications in Statistics-Theory and Methods*, 20(10), 3325-3340.

56. Ray. S.K. and Singh, R.K. (1981). Difference-cum- ratio type estimators, *J. Ind. Statist. Assoc*., 19, 147-151.

57. Robson, D.S. (1957). Applications of multivariate polykays to the theory of unbiased ratio-type estimation. *J. Amer. Statist. Assoc.*, 52, 511-522.

58. Roy, D.C. (2003). A regression-type estimator in two-phase sampling using two auxiliary variables. *Pak. J. Statist.*, 19(3), 281-290.

59. Ryu, J.B., Kim, J.M., Heo, T.Y. and Park, C.G. (2005). On stratified randomized response sampling. *Model Assisted Statistics and Applications*, 1(1), 31-36.

60. Saha, A. (2007). A simple randomized response technique in complex surveys. *Metron*, 65(1), 59-66.

61. Saha, A. (2008). A randomized response technique for quantitative data under unequal probability sampling. *Journal of Statistical Theory and Practice*, 2(4), 589-596.

62. Sahoo, J., Sahoo, L.N. and Mohanty, S. (1994). An alternative approach to estimation in two-phase sampling using two auxiliary variables. *Biom. J.*, 36, 293-298.

63. Samiuddin, M. and Hanif, M. (2007). Estimation of population mean in single and two phase sampling with or without additional information. *Pak. J. Statist.*, 23(2), 99-118.

64. Sanaullah, A., Hanif, M. and Asghar, A. (2016). Generalized Exponential Estimators for Population Variance Under Two-Phase Sampling. *International Journal of Applied and Computational Mathematics*, 2(1), 75-84.

65. Searls, D.T. (1964). The utilization of a known coefficient of variation in the estimation procedure. *J. Amer. Statist. Assoc.*, 59, 1225-1226.

66. Sharma, P., Verma, H.K., Sanaullah, A. and Singh, R. (2013). Some exponential ratio-product type estimators using information on auxiliary attributes under second order approximation. *International Journal of Statistics & Economics* 12(3), 58-66.

67. Singh, B.K. and Choudhury, S. (2012). Exponential chain ratio and product type estimators for finite population mean under double sampling scheme. *Journal of Science Frontier Research in Mathematics and Design Sciences*, 12(6), 0975-5896.

68. Singh, G.N. and Upadhyaya, L.N. (1995). A class of modified chain type estimators using two auxiliary variables in two-phase sampling. *Metron*, 53(3-4), 117-125.

69. Singh, H.P. and Solanki, R.S. (2013). A new procedure for variance estimation in simple random sampling using auxiliary information. *Statistical Papers*, 54(2), 479-497.

70. Singh, H.P. and Vishwakarma, G.K. (2007). Modified exponential ratio and product estimators for finite population mean in double sampling. *Austrian Journal of Statistics*, 36(3), 217-225.

71. Singh, H.P. and Espejo, M.R. (2000). An improved class of chain regression estimators in two phase sampling. *Statistics and Decision*, 18, 205-218.

72. Singh, H.P. and Espejo, M.R. (2003). On linear regression and ratio–product estimation of a finite population mean. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52(1), 59-67.

73. Singh, H.P. and Gangele, R.K. (1999). Classes of almost unbiased ratio and product type estimators in two phase sampling. *Statistica*, 59(1), 109-124.

74. Singh, H.P. and Gorey, S.M. (2016). Efficient Estimation of population Mean of Sensitive Variable in Presence of Scrambled Responses. *Communications in Statistics-Theory and Methods*, (accepted).

75. Singh, H.P., Singh, A.K. and Solanki, R.S. (2014). Estimation of finite population variance using auxiliary information in sample surveys. *Statistica*, 74(1), 99-116.

76. Singh, H.P., Tailor, R., Singh, S. and Kim, J.M. (2008). A modified estimator of population mean using power transformation. *Statistical Papers*, 49(1), 37-58.

77. Singh, H. and Tarray, T. (2014). An Improved Randomized Response Additive Model. *Sri Lankan Journal of Applied Statistics*, 15(2), 131-138.

78. Singh, R. and Malik, S. (2014). Improved estimation of population variance using information on auxiliary attribute in simple random sampling. *Applied Mathematics and Computation*, 235, 43-49.

79. Singh, R., Chauhan, P., Sawan, N. and Smarandache, F. (2009). *Improved exponential estimator for population variance using two auxiliary variables*. *arXiv preprint arXiv:0902.0126*.

80. Singh, S. (2010). Proposed optimal orthogonal new additive model (POONAM). *Statistica*, 70(1), 73-81.

81. Singh, S., Joarder, A.H. and King, M.L. (1996). Regression analysis using scrambled responses. *Australian & New Zealand Journal of Statistics*, 38(2), 201-211.

82. Singh, S., Sedory, S.A. and Arnab, R. (2015). Estimation of finite population variance using scrambled responses in the presence of auxiliary information. *Communications in Statistics-Simulation and Computation*, 44(4), 1050-1065.

83. Singh, V.K. and Singh, G.N. (1991). Chain type regression estimators with two auxiliary variables under double sampling scheme. *Metron*, 49(1-4), 279-289.

84. Singh, V.K., Singh, H.P. and Singh, H.P. (1994). Estimation of ratio and product of two finite population means in two-phase sampling. *Journal of Statistical Planning and Inference*, 41(2), 163-171.

85. Solanki, R.S., Singh, H.P. and Rathour, A. (2012). An alternative estimator for estimating the finite population mean using auxiliary information in sample surveys. *ISRN Probability and Statistics, 2012*.

86. Sousa, R., Gupta, S., Shabbir, J. and Corte-Real, P. (2014). Improved mean estimation of a sensitive variable using auxiliary information in stratified sampling. *Journal of Statistics and Management Systems*, 17(5-6), 503-518.

87. Sousa, R., Shabbir, J., Real, P.C. and Gupta, S. (2010). Ratio estimation of the mean of a sensitive variable in the presence of auxiliary information. *Journal of Statistical Theory and Practice*, 4(3), 495-507.

88. Srivastava, S.K. (1967). An estimator using auxiliary information in sample surveys. *Calcutta Statistical Association Bulletin*, *16*(2-3), 121-132.

89. Srivastava, S.K. (1970). A Two-Phase Sampling Estimator in Sample Surveys. *Australian & New Zealand Journal of Statistics*, 12(1), 23-27.

90. Strachan, R., King, M. and Singh, S. (1998). Theory and Methods: Likelihood-based Estimation of the Regression Model with Scrambled Responses. *Australian & New Zealand Journal of Statistics*, 40(3), 279-290.

91. Tarray, T.A. and Singh, H.P. (2015). Some improved additive randomized response models utilizing higher order moments ratios of scrambling variable. *Model Assisted Statistics and Applications*, 10(4), 361-383.

92. Tracy, D.S. and Singh, H.P. (1997). A family of estimators for population mean using prior/auxiliary information in sample surveys. *Jour. Appl. Statist. Sci.*, 6(1), 1-20.

93. Tracy, D.S. and Singh, H.P. (1999). A general class of chain regression estimators in two-phase sampling. *Journal of Applied Statistical Science*, 8(4), 205-216.

94. Upadhyaya, L.N., Singh, H.P., Chatterjee, S. and Yadav, R. (2011). Improved ratio and product exponential type estimators. *Journal of Statistical Theory and Practice*, 5(2), 285-302.

95. Walsh, J.E. (1970). Generalization of ratio estimate for population total. *Sankhya A*, 32, 99-106.

96. Warner, S.L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309), 63-69.

97. Warner, S.L. (1971). The linear randomized response model. *Journal of the American Statistical Association*, 66(336), 884-888.

98. Watson, D.J. (1937). The estimation of lead areas. *J. Agri. Sci.*, 27, 475.

99. Wilson, E.B. (1927). Probable inference, the law of succession and statistical inference. *Journal of the American Statistical Association*, 22(158), 209-212.

100. Yadav, S.K. and Kadilar, C. (2013). Improved Exponential Type Ratio Estimator of Population Variance. *Revista Colombiana de Estadística*, 36(1), 145-152.

101. Yadav, S.K., Mishra, S.S., Shukla, A.K. and Tiwari, V. (2015). Improvement of Estimator for Population Variance using Correlation Coefficient and Quartiles of The Auxiliary Variable. *Journal of Statistics Applications & Probability*, 4(2), 259-263.

102. Yadav, S.K., Misra, S. and Mishra, S.S. (2016). Efficient Estimator for Population Variance Using Auxiliary Variable. *American Journal of Operational Research*, 6(1), 9-15.

103. Yan, Z., Wang, J. and Lai, J. (2008). An efficiency and protection degree-based comparison among the quantitative randomized response strategies. *Communications in Statistics—Theory and Methods*, 38(3), 400-408.